Understanding Photostability of Biomolecules using Multi-reference Quantum Chemical Methods

A thesis submitted towards partial fulfilment of
PhD Programme

by

AVDHOOT DATAR

DEPARTMENT OF CHEMISTRY

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

For my mother, Vandana Datar

Declaration

I declare that the matter embodied in the thesis entitled "Understanding Photostability of Biomolecules using Multi-reference Quantum Chemical Methods" represents my idea in my own words and where others ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

A.S. Datar

(Mr. Avdhoot Datar)
Reg. No. 20123169

Pune, India
15th October, 2018

Acknowledgements

My PhD has been a very rewarding experience throughout. First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Anirban Hazra. He has helped me grow as an independent researcher by allowing me to choose and develop my research problem. He has always been very appreciative of an individual's interest and willingness to learn, and has never perceived a lack of knowledge as an obstacle. He has also inculcated in me certain values, which I will always carry with me. For instance, if you work hard enough on a problem, it is bound to be solved. I have learned from him the value of doing things systematically, taking care with every piece of work, and always striving to achieve perfection, irrespective of the task at hand.

I would like to thank Dr. Arnab Mukherjee, for all the discussions we have had. There were several occasions, when I felt that the problem at hand was unsolvable, but then seemed simple after my discussions with him. I also thank the members of my Research Advisory Committee: Dr. Prasenjit Ghosh, and Dr. Kumar Vanka, for their valuable suggestions and spurring me on to complete my thesis.

I am grateful to IISER Pune not just for my fellowship, but also the infrastructure which has eliminated so many inconveniences.

I especially would like to thank Dr. Debashree Ghosh, Dr. Arijit Bhattacharyay, Dr. Pankaj Mandal, Dr. Sabyasachi Pal, and Prof. Rajiv K. Pathak, all of whom have given me their valuable time at some point or the other to talk about the problems I was facing in my research. I wish to thank my collaborators, Dr. Sayan Bagchi, Dr. Jeetender Chugh, and Mr. Pranab Deb for spectroscopic experiments. I also thank Dr. Pinaki Talukdar for having faith in me, by giving me the opportunity to perform computational studies to complement his experiments.

It would be remiss to not talk about the contributions of my colleagues towards my PhD. Mahesh Gudem has always answered any question I had, irrespective of how trivial. His help was invaluable to me when I shifted from the 'Independent Electron Surface Hopping' method and started working on electronic structure problems. Meghna Manae was always there whenever any document needed editing or I needed a sounding board to check if I was on the right track. I also thank her for her unwavering support. I thank my juniors Khushboo and Aditya; teaching them instilled me with confidence regarding my own abilities.

I would like to thank Harshad Paithankar for all the input from an Organic chemist's perspective, concerning my understanding of Barbituric acid. I also thank him for per-

forming NMR experiments on BA. But mostly, I thank him for all the times he brewed tea in his room at the end of a very long day. I thank Reman Singh who taught me how to perform molecular dynamics simulations with immense patience. But more than that, I thank him for being such a great friend and an unending source of optimism.

I thank Satyam Srivastava, for all the involved discussions on electronic structure theory, and his constant encouragement. I thank Mayukh Kansari for his extremely large grin, which always made me smile. I thank Bappa Ghosh and Divya Singh for such a vibrant atmosphere in the lab, Jyotirmoy Roy for making conferences such fun, Abhishek Swarnakar for all those times that he was my alarm clock, Neha for the late night discussions on transient absorption spectroscopy, and Niharika, Subbu and Nandha for all their help with using Quantum Espresso. I also thank my longtime friends, Mangesh Thore, Prasad Bagul, and Kiran Limaye, for simply being there whenever I needed them.

I gratefully acknowledge the financial support extended to me by Hrushikesh Joshi, Pravinji, and Dharap Kaka, without which I would have found it hard to continue with my education.

My family has always been very supportive of my choices. In this regard, I especially thank Aai for her unflinching support and patience. I thank Kaka, Mausi, and my in-laws for their encouragement. Finally, I thank Varada: this journey would not have been possible without you by my side.

Abstract

Photostability is the property of a molecule to not chemically react when excited with light. This is a prerequisite for the origin and preservation of life since the ambient environment has ample radiation from the sun, which can be potentially damaging. Here, we have studied the photodecay mechanisms of two biologically relevant molecules: 5,6-dihydroxyindole (DHI) and barbituric acid (BA). Our approach involves obtaining pathways on the excited state potential energy surfaces for deactivation, which usually involve bond stretched geometries and points of degeneracy like conical intersections (CIs). Such calculations require multi-reference electronic structure methods and accordingly we have used the complete active space self-consistent field (CASSCF) approach and its extensions. Unlike other commonly used electronic structure methods, these require significant user input based on chemical intuition for an accurate description of excited state processes.

DHI is a monomeric unit of eumelanin, a natural sunscreen present in human skin. Assuming that DHI photophysics will be reflective of eumelanin photophysics, a bottom-up approach has been employed to study its photophysics. Experiments on DHI show excitation energy dependent fluorescence kinetics and our study reveals the reason for this. We found two planar and one non-planar CIs through which photodecay of the excited state can take place. Accessing these CIs require the molecule to overcome energy barriers, which is possible when the initial excitation is to higher vibrational states of the excited electronic states. Thus, higher energy leads to shorter excited state lifetimes. This result is significantly different from earlier results obtained with single-reference methods. Our study also shows that isolated monomeric DHI cannot explain the ultrafast deactivation of eumelanin.

BA is proposed to be a prebiotic nucleobase which played a role in the emergence of

RNA-based life. For BA to be a candidate for proto-RNA, it must have been photostable to survive the intense UV radiation present at the time when the ozone layer was absent. BA can exist as two tautomers: keto and enol. Theoretical calculations on isolated BA predict that the keto tautomer is more stable than enol. However, recent reports have shown that crystalline BA exists in the enol form. We have explored the relative stability of this molecule in the aqueous environment. Our quantum calculations on BA-water clusters, where the clusters were chosen systematically using a molecular dynamics based scheme, show that the enol tautomer gets stabilized more than the keto tautomer due to solvation.

Recently, transient absorption experiments on BA have explored its photodecay dynamics. Our calculated photodecay pathways of BA help in interpreting these experiments. Further, we also compare the photodecay dynamics of BA with that of Uracil, a canonical RNA nucleobase, and find that despite strong structural similarities, the two molecules follow different photodecay pathways.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Molecules on absorbing UV-visible radiation, typically get electronically excited. Such electronically excited molecules can potentially undergo chemical reactions due to the excess energy available to them. Photostable molecules, on the other hand, possess the property to convert this excess energy to heat in the form of molecular vibrations, without undergoing any chemical change.

A more detailed picture of the processes following electronic excitation is shown in Figure 1.1. After electronic excitation, from a singlet ground state, the molecule starts relaxing toward the minimum energy geometry on that excited state. From this geometry it can undergo radiative transition to the ground state - this process is called fluorescence. While approaching the minimum, the molecule can also find a region where there is a degeneracy with another excited state and internal conversion to this other excited state can take place. It can also populate a triplet state, via a singlet-triplet crossing, which is known as intersystem crossing. After reaching the triplet state, the molecule can either radiatively decay to the ground state, a process called phosphorescence, or cross back to the singlet manifold through a different singlet-triplet crossing point. Typically, the timescale of fluorescence is much faster than phosphorescence. The processes starting from the absorption of a photon till the point when the molecule decays back to the

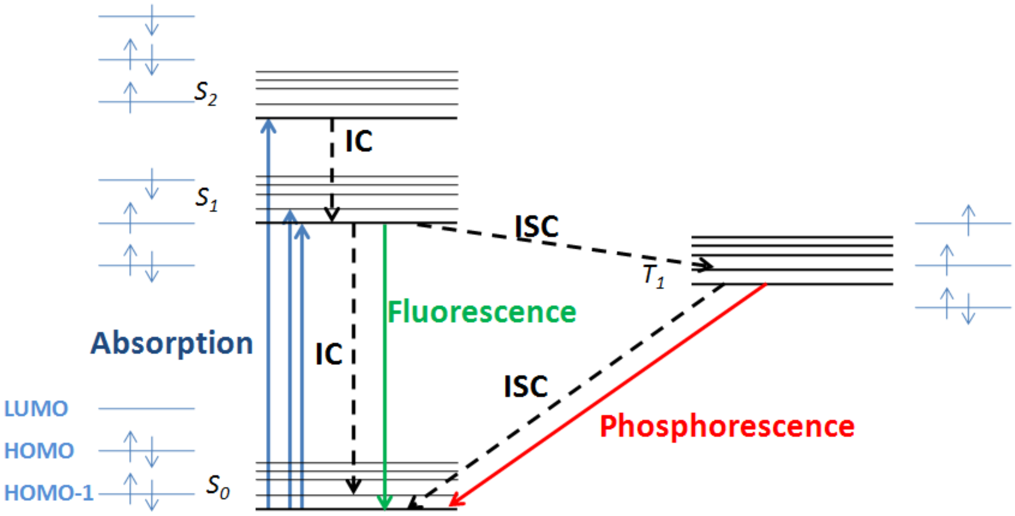


Figure 1.1: Jabolanski diagram

same ground state are called photophysical processes. These processes can be described by Jablonski diagrams.[1][2]

Biomolecules, such as melanin, DNA nucleobases, some amino acids, show a high degree of photostability towards UV radiations. In this context, biomolecules can be divided into two classes. An example of the first class of molecules are melanins: molecules that act as a filter toward UV radiations. By converting these radiations into heat, they protect skin and internal tissues from harmful radiations.[3] The second category consists of molecules whose primary function is not photostability upon UV excitation, however, photostability is an important requirement for their general functionality. Prominent examples of such molecules are DNA and RNA nucleobases - uracil, thymine, cytosine, adenine, and guanine, which are information carriers of the genetic code. After absorbing UV radiations, these molecules efficiently undergo internal conversion on an ultrafast timescale and are extremely photostable.[4] Extensive examples highlight the importance of photostability in biomolecules both for the protection and preservation of life. In this thesis we present our studies, regarding mechanistic understanding of photostability of

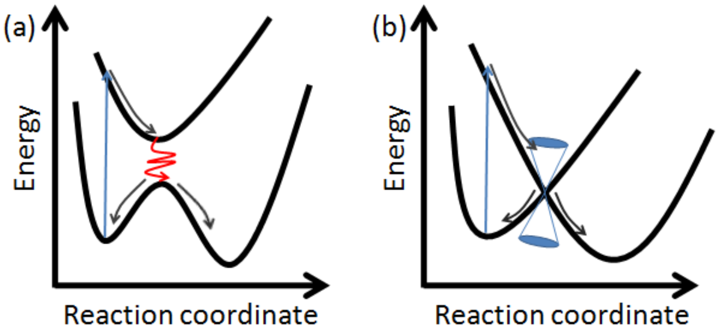


Figure 1.2: Avoided crossing and conical intersection.

a building block of eumelanin and a proto-RNA nucleobase.

Eumelanin belongs to a class of biopolymers known as melanins. It is present in the skin, eyes and hair. It absorbs strongly in the UV region of the electromagnetic spectrum as well as possesses close to unity quantum yield for nonradiative decay. Therefore, it is believed that eumelanin plays an important role in UV protection.[3] As eumelanin acts as a natural sunscreen, understanding the mechanisms of its photostability would be useful in translating this knowledge in the development of artificial sunscreens to prevent damages due to long UV exposure. Recently, photostability of similar molecules have been studied for understanding their potential role in sunscreens.[5][6]

Contemporary RNA and DNA nucleobases are the product of evolution. It is believed that on prebiotic earth, the role of information transfer would have been played by some other heterocycles, and not the DNA bases we have today. Further, the ozone layer was not present to absorb UV radiations around the earth in prebiotic time.[7][8] Therefore, the proto-RNA, from which current RNA has evolved, was exposed to a high flux of UV radiations. Photoprotection against UV radiation therefore, must have played a decisive role in selecting the nucleobases of proto-RNA.[9] Impact of photostability on the evolution of RNA and specific choices of the current nucleobases, can be revealed by studying the photostability of proposed proto-RNA molecules.

In this thesis, we have used computational approach for studying photostability. Computational chemistry can provide mechanistic insight into photostability, *i.e.*, one can find the role of twisting or stretching of certain bonds which governs the nonradiative decay. The potential energy surface (PES), which is a function of energy of electronic states along various molecular geometries, is a very useful concept in studying mechanisms responsible for nonradiative decays. There are two approaches of performing computational studies to understand nonradiative decay processes: static and dynamical. In static calculations, PESs of relevant electronically excited and ground states are examined for finding the pathways which lead to a transition from the excited states to the ground state. Dynamical calculations, on the other hand, are performed by studying nuclear motions on the PESs constructed from static calculations. Given the computational cost of dynamical calculations, as a first step for understanding various excited state processes, static calculations are often performed with accurate electronic structure methods. From these calculations important internal coordinates, which take part in the nonradiative decay, can be identified. Then dynamical calculations can be performed for the reduced space of these important internal coordinates. The results obtained with static calculations are presented, in this thesis.

A photoexcited molecular system can switch from one electronic state to another electronic state when they are close in energy. *Avoided crossings*, as shown in Figure 1.2 a, are often used to explain such a transfer. An avoided crossing can be described as a region of the PES where two electronic states are very close in energy. However, due to the presence of a finite gap in electronic energy, nonradiative decays that take place on ultrafast timescales cannot be explained with avoided crossings. Zimmerman[10][11], Michl[12], and Teller[13] independently introduced the idea of *photochemical funnels* (Figure 1.2 b). It was suggested that nonradiative decay takes place from geometries where two electronic states become degenerate. These real crossing points are known as *conical intersections* (CI). CIs play a very important role in explaining photostability to molecules. CIs can be considered to have a similar importance in excited-state processes, as transition states have in ground state chemistry.[14]

For the correct description of the PES near a CI, one has to use multi-reference electronic structure methods in the calculation. Unlike, single-reference methods, like HF, DFT, CCSD etc., these methods can simultaneously describe multiple electronic states. In our studies, we have used the CASSCF and CASPT2 methods, which are described in Chapter 2 of this thesis.

Outline of thesis

In Chapter 2, we briefly describe the computational methods used in this study. For studying ground state reactions, the electronic structure methods employed are MP2, CCSD, DFT. However, different theoretical methods are necessary for studying photoexcited molecular reactions, as these studies involve situations of bond breaking, bond stretching, and crossing of electronic states. These methods are known as multi-reference methods. A brief introduction to CASSCF and CASPT2 methods is provided.

This thesis includes our work on understanding mechanisms responsible for photostability of 5,6-dihydroxyindole (DHI) and barbituric acid (BA). DHI is a building block for melanin which protects life from UV radiations, while BA is a proto-RNA nucleobase. This study will be useful towards building a understanding of photostability of complex biomolecules.

Chapter 3 describes our study on mechanisms for nonradiative decay of DHI. DHI is a monomeric unit of eumelanin, a natural sunscreen present in human skin. Assuming that DHI photophysics will be reflective of eumelanin photophysics, a bottom-up approach has been employed to study its photophysics. Experiments on DHI show excitation energy dependent fluorescence kinetics and our study reveals the reason for this. We found two planar and one non-planar CIs through which photodecay of the excited state can take place. Accessing these CIs require the molecule to overcome energy barriers, which is possible when the initial excitation is to higher vibrational states of the excited electronic states. Thus, higher energy leads to shorter excited state lifetimes. This result is significantly different from earlier results obtained with single-reference methods. Our study also shows that isolated monomeric DHI cannot explain the ultrafast deactivation of eumelanin.

In Chapter 4, we provide our results on the role of water in stabilizing the enol tautomer of BA. BA is proposed to be a prebiotic nucleobase which played a role in the emergence of RNA-based life. BA can exist as two tautomers: keto and enol. Theoretical calculations on isolated BA predict that the keto tautomer is more stable than enol. However, recent reports have shown that crystalline BA exists in the enol form. We have explored the relative stability of this molecule in the aqueous environment. Our quantum calculations on BA-water clusters, where the clusters were chosen systematically using a molecular dynamics based scheme, show that the enol tautomer gets stabilized more than the keto tautomer due to solvation.

For BA to be a candidate for proto-RNA, it must have been photostable to survive the intense UV radiation present at the time when the ozone layer was absent. Recently, transient absorption experiments on BA have explored its photodecay dynamics. Our calculated photodecay pathways of BA help in interpreting these experiments. These results are discussed in Chapter 5.

Chapter 2

Theoretical Background and Methods

This chapter has two parts. First, we discuss the well-known idea of potential energy surfaces (PES), which is a useful tool for describing mechanistic picture of photostability. We also provide an introduction to the concept of conical intersections (CIs), which play an important role in nonradiative decay of molecules. In the second part, we briefly describe electronic structure methods, that are used for performing the studies presented in this thesis. A more detailed description of these methods can be found in Ref [15].

2.1 Describing electronic and nuclear motion

The idea of PES provides an understanding of the nature of electronic states at various molecular geometries. It emerges from the Born-Oppenheimer (BO) approximation.[16] The BO approximation forms the basis for computational chemistry. Within this approximation one can study nuclear and electronic motions separately.

The total Hamiltonian for molecular system containing N electrons and M atoms can be written as:

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{eN}(\mathbf{r}, \mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) \tag{2.1}$$

In this Hamiltonian, the first two terms represent the kinetic energies of electrons and nuclei respectively. The third term in the equation accounts for electrostatic attraction between electrons and nuclei. The last two terms are for nuclear-nuclear and electronic-electronic repulsion terms, respectively. $\mathbf{r}$ and $\mathbf{R}$ denote electronic and nuclear coordinates, respectively. Without the term $\hat{V}_{eN}(\mathbf{r}, \mathbf{R})$ electronic and nuclear Hamiltonian could be written separately. This separation would facilitate the description of the total wavefunction of the system as a product of the electronic and nuclear wavefunctions. Contribution of $\hat{V}_{eN}(\mathbf{r}, \mathbf{R})$ is not small and thus it cannot be neglected. BO approximation provides a solution to decouple the electronic and nuclear motion without neglecting this term. As the mass of an electron is ~2000 times smaller than the mass of a proton (lightest nucleus of Hydrogen atom), nuclei move much more slowly compared to electrons. BO approximation assumes that electrons adjust their motion instantaneously as the nuclear geometry is changed. In other words, within this approximation we write the electron-nuclear coupling term as $\hat{V}_{eN}(\mathbf{r}; \mathbf{R})$, i.e., a parametric dependence of nuclear coordinates ($\mathbf{R}$) is assumed. For a fixed nuclear geometry, we solve time independent Schrödinger equation:

$$\hat{H}_{el}\Psi_{el}(\mathbf{r}; \mathbf{R}) = E_{el}(\mathbf{R})\Psi_{el}(\mathbf{r}; \mathbf{R}) \tag{2.2}$$

where, $\Psi_{el}(\mathbf{r}; \mathbf{R})$ is the electronic wavefunction and $E_{el}(\mathbf{R})$ is the electronic energy at a fixed nuclear geometry $\mathbf{R}$. In computational chemistry, one of the oft repeated calculations is solving this electronic Schrödinger equation at various nuclear geometries. To solve the nuclear Schroödinger equation this electronic energy (and nuclear repulsion energy) serves as the potential energy. It can be imagined that nuclear motion takes place on the PES. The approximation aspect of BO approximation lies in ignoring the coupling between two electronic states due to change in nuclear geometry, *i.e.*,

$\langle \Psi_i(\mathbf{r};\mathbf{R})| \nabla_A |\Psi_j(\mathbf{r};\mathbf{R})\rangle$ and $\langle \Psi_i(\mathbf{r};\mathbf{R})| \nabla_A^2 |\Psi_j(\mathbf{r};\mathbf{R})\rangle$, for all electronic states i and j. While studying excited states, it is possible that two or more electronic states lie close in energy.

Conical Intersections

The nonadiabatic coupling between two electronic states is inversly proportional to the difference in energy of two electronic states. Two or more electronic states having a small energy difference are called coupled electronic states. BO approximation fails when two or more electronic states are coupled, as the nonadiabatic coupling terms in such cases cannot be ignored. While studying photochemistry one has to often find molecular geometries having degenerate electronic states. As stated earlier, CIs are geometries that act as photochemical funnels for the transfer of molecular population from one electronic state to another electronic state. Describing these regions of the PES is central to studying nonradiative decay mechanisms.

Consider two interacting electronic states ϕ_1 and ϕ_2. The electronic Hamiltonian in the basis of these two electronic states is written as follows:

$$H_{el}(\mathbf{R}) = \begin{bmatrix} H_{11}(\mathbf{R}) & H_{12}(\mathbf{R}) \\ H_{21}(\mathbf{R}) & H_{22}(\mathbf{R}) \end{bmatrix} \tag{2.3}$$

here, $H_{ij}(\mathbf{R}) = \langle \phi_i(\mathbf{r};\mathbf{R})| \hat{H}_{el}(\mathbf{r};\mathbf{R}) |\phi_j(\mathbf{r};\mathbf{R})\rangle$. For this matrix Hamiltonian to give degenerate electronic states, it must follow the conditions given below.

$$\begin{aligned} H_{11}(\mathbf{R}) &= H_{22}(\mathbf{R}) \\ H_{12}(\mathbf{R}) &= H_{21}(\mathbf{R}) = 0 \end{aligned} \tag{2.4}$$

These two conditions mean that there are at least two nuclear degrees of freedom along which the degeneracy of the electronic states can be broken in first order. These two directions are necessary to form the 'cone shape' of conical intersections in nuclear space. In diatomic molecules only one degree of freedom (bond length) exists, therefore two states of same (spatial and spin) symmetry can not cross. This is know as the *n*oncrossing rule formulated by von Neumann and Wigner in 1929.[17] As several nuclear degrees of freedom present in polyatomic molecules, it is possible in these systems that two states of same symmetry can have an intersection.

Denoting these two degrees of freedom as $\mathbf{x}_1$ and $\mathbf{x}_2$. Within the linear approximation, we can show that these vectors are gradient difference vector and nonadiabatic coupling vector,

$$\begin{aligned} \mathbf{x}_1 &= \frac{\partial(E_1 - E_2)}{\partial \mathbf{R}} \\ \mathbf{x}_2 &= \left\langle \phi_1 \middle| \frac{\partial H_{12}}{\partial \mathbf{R}} \middle| \phi_2 \right\rangle \end{aligned} \tag{2.5}$$

Here, E_1 and E_2 are adiabatic energies obtained by diagonalizing the Hamiltonian matrix. A plane formed by $\mathbf{x}_1$ and $\mathbf{x}_2$ is known as *branching plane*. Degeneracy of the states is lifted in the first order (linearly), while moving away from the CI in the branching plane. As two PESs of interacting electronic states are crossing in the branching plane, the topology of the PESs in this plane looks like a double cone. It is very important to note that the branching plane can be uniquely defined by gradient difference vector and derivative coupling vector, and hence CIs are not single points in geometric space but there are infinitely many CIs along the *intersection seam*, which are orthogonal to the branching plane. For a molecule with M internal degrees of freedom (for a N-atomic molecule, M=3N-6), there are M-2 coordinates orthogonal to the branching plane along which degeneracy is maintained and such a hyperline in geometric space is called as the intersection seam.[18][19] CIs are high dimensional entities, and are thus more accessible in a photochemical reaction. The most relevant CI in a photochemical reaction however, is the minimum energy geometry on the intersection seam. It is called the minimum energy conical intersection (MECI). Generally in photochemical studies CI denotes MECI, due to its relevance in photochemical reactions. Modern quantum chemistry codes routinely optimize CIs to find MECI. Here onward the term CI is used to represent MECI, unless explicitly stated otherwise.

Pathway for Nonradiative Decay

Mechanistic studies of photostability involves description of pathways leading to nonradiative decay of excited states. These pathways include minimum energy geometries, transition states, and crossing electronic states (e.g. CIs and singlet-triplet crossing points). Describing topology of electronic states along multiple degrees of freedom is an arduous task and comprehending the mechanism responsible for nonradiative decay from multidimensional PESs is not feasible. Instead, comparison of CI geometry with respect to Franck-Condon geometry helps in understanding possible internal coordinates which take part in the nonradiative decay. Depending on the complexity of internal coordinates involved, there two possible ways of describing potential energy curves (PEC) that lead to nonradiative decay. PEC is differentiated from a PES in that it is a one-dimensional representation, as opposed to a PES which is M-dimensional, and hence computationally unfeasible.

Relaxed Scan

Relaxed scan of PEC involves constraining one particular internal coordinate and optimizing the geometry for all other internal coordinates of the molecule. In this fashion a one-dimensional PEC is plotted to investigate the nature of electronic states. While studying nonradiative decay, which is governed by a significant distortion of a single internal coordinate, relaxed scan approach is employed. By comparison of CI geometry with the FC geometry it can be examined, if there is an internal coordinate which is changing significantly. For example, $\pi\sigma^*/S_0$ CI, in heteroaromatic systems often involves significant elongation of X-H (X: heteroatom N,O) bond length.[20] For studying nonradiative decay via $\pi\sigma^*/S_0$ CI often relaxed scan approach is applied along X-H bond stretching coordinate.

Performing relaxed scan is computationally expensive, as one has to do numerous geometry optimizations along the constrained reaction coordinate. However, relaxed scan of PEC provides an accurate description of barriers involved in accessing CI from the FC geometry. This approach is not applicable if CI geometry has distortions along multiple internal coordinates with respect to FC geometry.

Linearly Interpolated Internal Coordinates

In this approach intermediate geometries are generated by linear interpolation of FC geometry to the critical geometry (CI, minimum, or intersystem crossing) on the excited state taking part in nonradiative decay. Then, energy profiles for relevant excited states and ground state is calculated at the geometries generated. The pathway obtained by this approach may not be the actual pathway followed by molecules in nonradiative decay, as gradients of excited states are not taken into account. However, barriers calculated for accessing critical geometries from the FC geometry are upper bounds for the actual barriers experienced by the molecule.

This approach is computationally less demanding, as there is no geometry optimization involved. This approach is widely applied for explaining mechanisms of nonradiative decay of DNA-nucleobases which are known to have out-of plane motions for accessing $\pi\pi^*/S_0$ or $n\pi^*/S_0$ CIs.[21]

So far we have discussed the idea of PES and various ways of representing the pathways leading to nonradiative decay via CI. Now we introduce the electronic structure methods that are routinely used for constructing these pathways.

2.2 Single-Reference Methods

Finding the solutions to the electronic Schrödinger equation with a multi-electron (more than 1) system is a central problem in electronic structure theory. Hatree-Fock approximation is conceptually the most fundamental approach for solving this problem. This is an independent particle approximation, i.e., the electronic wavefunction is assumed to be a product of one electron wavefunctions. To fulfill Pauli's exclusion principle, this product is written in terms of a determinant, which is called as Slater determinant. Slater determinant for n-electrons is given below.

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1)\phi_2(1)...\phi_n(1) \\ \phi_1(2)\phi_2(2)...\phi_n(2) \\ ...... \\ \phi_1(n)\phi_2(n)...\phi_n(n) \end{vmatrix} \tag{2.6}$$

here, each ϕ_i is a one electron wavefunction. These wavefunctions are written as linear combinations of atomic orbitals as, $\phi_i = \sum_\alpha c_\alpha \xi_\alpha$. Here, ξ_α are atomic orbitals. Then the linear variation method is applied, for finding the Slater determinant with minimum energy, i.e., $E_0 = \langle \Psi | \hat{H}_{el} | \Psi \rangle$. This leads to n eigenvalue equations as shown below.

$$f(i)\phi_a(i) = \epsilon_a \phi_a(i) \tag{2.7}$$

Here, ϵ_a is the eigenvalue of orbital ϕ_a. $f(i)$ is the Fock operator defined as:

$$f(i) = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{A=1}^{M} \frac{1}{4\pi\epsilon_0} \frac{Z_A e^2}{|\mathbf{r}_i - \mathbf{R}_A} + \nu^{HF}(i) \tag{2.8}$$

$\nu^{HF}(i)$ is the effective energy, so-called mean field, felt by the i^{th} electron due to presence of other electrons. For calculating $f(i)$ we need to find orbitals for all the electrons, including the i^{th} electron. By solving the eigenvalue equations we find a set of orbitals

$\{\phi_a\}$. HF is a self-consistent field approach for finding the mean field experienced by electrons.

HF method does not account for both static and dynamical correlation, which is experienced due motion of electrons. Dynamical correlation is added in post-HF methods, for example MP2 [22], CCSD [23][24], and others, by constructing multi-electron wavefunction as a linear combination of Slater determinants. The linear combination of Slater determinant is constructed in such a way that it is an eigenfunctions of the spin operator ($\hat{S}^2$), and this is called a configuration state function (CSF).

The Configuration Interaction is the simplest of the post-HF methods. In this method, the electronic wavefunction is constructed as a linear combination of Slater determinants or CSFs, as shown below.

$$\Psi = C_0\Psi_0 + \sum_s C_s\Psi_s + \sum_d C_d\Psi_d + \sum_t C_t\Psi_t \quad ... \tag{2.9}$$

here, Ψ_0 is the Slater determinant obtained from HF calculation, Ψ_s are all the Slater determinants obtained by singly replacing an occupied HF orbital with each unoccupied orbital obtained, to represent a singlet excitation. Similarly d represents double replacement or double excitation and so on. Considering all possible excited determinants leads to a full Configuration Interaction approach. For a given basis, full Configuration Interaction gives exact results, but this approach is not computationally feasible.

Electronic structure methods based on the HF method, called as single-reference methods, are not suitable for describing molecular systems with degenerate-ground state and while dealing with bond breaking situations. In the excited state, where two or more electronic states often cross each other due smaller geometrical distortions, the single-reference methods do give accurate results. In such situations one has to use multi-configurational methods.

2.3 Multi-reference methods

Complete-Active Space Self-Consistent Field (CASSCF) method

CASSCF is a multi-reference method.[25][26] Similar to Configuration Interaction, in this method the wavefunction is expanded as a linear combination of CSFs. The molecular orbital space is divided into three categories: the inactive orbitals, the active orbitals, and the secondary orbitals. The inactive orbitals remain doubly occupied, while secondary orbitals are kept unoccupied, during the course of the calculation. On the other hand active orbitals, which contain some occupied and some unoccupied orbitals obtained from HF calculation, have partial occupancy between 0 and 2 in a CASSCF calculation. The CASSCF wavefunction is generated by full Configuration Interaction type expansion within the active space, composed of active orbitals, N_a and active electrons, n_e. An active space used for a CASSCF calculation, n_e including number of active electrons and N number of active orbitals, is represented by CAS(n_e,N). Figure 2.1 shows the schematic representation of the active space CAS(4,5).

$$N_{CSF} = \frac{2S+1}{N+1} \binom{N+1}{\frac{n_e}{2} - S} \binom{N+1}{\frac{n_e}{2} + S + 1} \tag{2.10}$$

The measure of difference between the Configuration Interaction method and the CASSCF method is that, in CASSCF not only are the weights of CSFs optimized but also the molecular orbitals describing each CSF. For a system with spin-state S, and active space CAS(n_e,N), total number of CSFs generated for the CASSCF wavefunction is given by equation 2.10. The time taken for CASSCF calculation scales in a factorial manner with the size of active space. In principle, one can include all orbitals and electrons in the active space, but this is essentially performing a full Configuration Interaction calculation including optimization of molecular orbitals in each CSF. The choice of active space is the key step in a CASSCF calculation. On one hand, a bigger active space makes the problem computationally unfeasible and on the other hand,s smaller active spaces

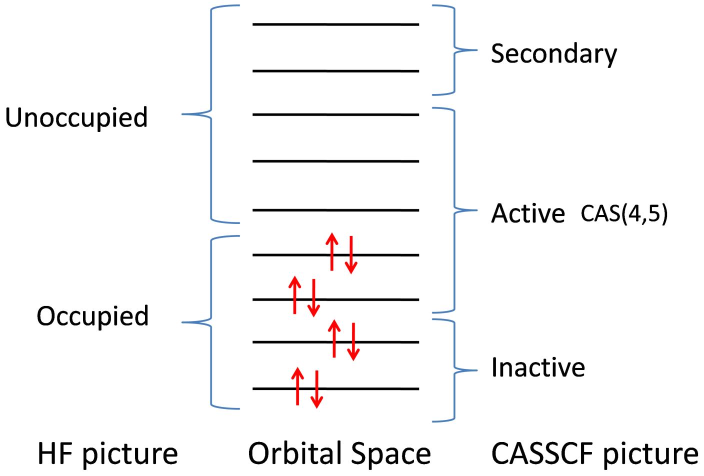


Figure 2.1: Description of the orbital space for the HF calculation and for the CASSCF calculation. The active space is shown of 4 electrons in 5 orbitals.

may not describe the chemical problem under study accurately. General guidelines for choosing an active space, while dealing problems involving excited states of heterocyclic systems are given below.

Selecting an active space

- For the calculation of vertical energies, if possible include all n, π and corresponding π^* orbitals.
- Calculate electronic excited states of the system using single-reference methods like EOM-CCSD, CC2, TDDFT, and others. This gives an idea of orbitals taking part in the excited states. Include all the important orbitals as seen from these calculations.
- Typically for heterocyclic systems, Rydberg-type orbitals located along the heteroatom (N, O, S..) take part in the excitation to lowest few excited states. This is not described correctly if the basis set does not contain diffuse-functions. First,

check if Rydberg-states are important for the photochemical reaction under study. For this purpose perform single-reference calculations with basis set including diffuse functions. If it turns out that Rydberg-states are important then one has to use the 'correct' basis set.

- If the reaction path involves stretching of any bond then σ and σ^* orbitals located along that bond should be included in the study.
- Occupied orbitals having occupation greater than 1.98 and unoccupied ones having less than 0.02, along the photoreaction path can be omitted from the active space.

While studying photostability of molecules, one has to study multiple electronic states simultaneously. Thus, we have use the state-averaged CASSCF method. In this method, the average energy of the selected electronic states is minimized. While CASSCF takes into account non-dynamical correlation, the lack of dynamical correlation is a serious drawback of the CASSCF method. To overcome this, one has to use methods like second-order perturbation correction to CASSCF (CASPT2), multi-reference configuration interaction (MRCI).[27] In our studies we have employed the CASPT2 method to include dynamical correlation.

CASPT2 method

As the name suggests the CASPT2 is a perturbative method with the zeroth order reference wavefunction is the CASSCF wavefunction.[28] In this thesis, we have employed two flavors of the CASPT2 method: single-state CASPT2 (SS-CASPT2)[29][30] and multi-state CASPT2 (MS-CASPT2) [31]. MS-CASPT2 method is computationally expensive. Thus, for calculating energies along reaction paths (where multiple calculations of energy are involved), we have used the SS-CASPT2 method.

Often CASPT2 calculations suffer from convergence issues due to intruder states. We have used a level-shift in the CASPT2 calculation to overcome this issue, as suggested in the literature.[32] Using level shifts in the calculation adds a constant to the zeroth order Hamiltonian, which can overcome the issue of intruder states. However, inclusion of level shift compromises the accuracy of the energy computation. It is advised that the level shift should not be higher than 0.3 a.u.

Due to high computational cost, CASPT2 is widely used for calculating single point energies of the system, while geometries are calculated with other methods like CASSCF. The commonly used protocol for studying reaction mechanisms in photophysics and photochemistry of molecules is CASPT2//CASSCF approach.[33] In this approach, geometries are optimized with the CASSCF method and then energy of the geometry is refined using the CASPT2 method. The CASPT2 and CASSCF calculations were performed with Molpro software package[34][35].

Chapter 3

Nonradiative Decay of 5,6-Dihydroxyindole

3.1 Introduction

Eumelanin is a pigment that imparts color to skin, hair, and eyes of humans and many other animals.[36] Chemically, eumelanin is a biopolymer that has a broad UV–visible absorption.[37][38] On photoexcitation, it undergoes nonradiative relaxation on the ultrafast time scale with quantum yields close to unity.[39] Because of these characteristic photoproperties, eumelanin plays a key role in protecting the skin from harmful UV radiation present in sunlight. In spite of its importance to life, its exact structure and its mechanisms for conversion of high frequency radiation into heat are not well-understood. Given its complex structure, a bottom-up approach is a pertinent strategy for mechanistic study of the photoproperties of eumelanin. This involves studying the photophysics and photochemistry of the building blocks or basic constituent molecules of eumelanin in great detail, using accurate quantum chemical methods or spectroscopic techniques, and then proceeding to understand the more complex structures formed from these constituents with less sophisticated approaches. The major building blocks of eu-

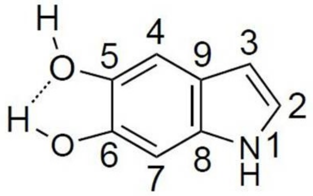


Figure 3.1: Structure of DHI in its stable form

melanin are the heterocyclic molecules 5,6-dihydroxyindole (DHI), indolequinone, and 5,6-dihydroxyindole carboxylic acid.[37][38][40] Herein, we present an in-depth study of the photophysics of DHI using multi-reference quantum chemical calculations.

The excited-state dynamics of DHI has been studied using transient absorption spectroscopy in 2009[41] and time-resolved fluorescence in 2011[42] by the research group of Sundström. These studies, performed in buffer solutions, found that the dissipative mechanisms after photoexcitation were strongly dependent on the excitation wavelength. On the basis of experimental evidence from the two studies, the following explanation was proposed for the wavelength-dependent dynamics: Shorter wavelengths populate the second optically bright state to an increasing extent. This state undergoes internal conversion to the lowest optically bright state (S_1) and can additionally undergo fast deactivation by electron transfer to the solvent, resulting in formation of a DHI radical cation and a solvated electron. The associated fluorescence lifetime is 110 ps. Longer-wavelength light leads to excitation of the S_1 state, for which the ultrafast excited-state deactivation by electron transfer to the solvent is not accessible and shows a relatively long lifetime of 1.7 ns.

Quantum chemical calculations, in combination with experimental studies, can provide detailed mechanistic insight into the excited-state dynamics. There have been several theoretical studies on the structure and photoproperties of DHI,[43–51] most of which were published before the experimental studies by Sundström and co-workers. DHI can have different tautomeric forms of which the structure where the nitrogen is in the amine form and both the oxygens are in the enol form has been found to be the most stable[45] (Figure 3.1). Vertical excitation energies of DHI have been calculated using semiempirical methods,[43, 44] time-dependent density functional theory.[45, 47, 49, 50] The photophysics and photochemistry of DHI on excitation to the first bright state has been investigated by Sobolewski and Domcke using the CC2 (approximated singles and doubles coupled-cluster) method.[48] On absorbing light, the formation of a photoproduct via excited-state intramolecular proton transfer (ESIPT) from the hydroxyl group at position 5 to the carbon at position 4 is predicted (refer to Figure 3.1 for atom numbering). Further, the photoproperties of the ESIPT product 4-hydro-6-hydroxyindole-one (HHI) have been explored. The electronic structure approach used in these studies, CC2, has been used in a variety of excited-state applications and is similar in accuracy to MP2 for ground states. However, being a single reference method it fails in regions where two states are degenerate and is therefore not able to accurately describe potential energy surfaces at conical intersections (CIs). In the present study the deactivation mechanisms of DHI after photoexcitation to its first and second optically bright electronic states are investigated by optimizing relevant CIs and estimating barriers to access these CIs from the Franck-Condon (FC) region. We used multi-reference electronic structure approaches the complete active space self-consistent field (CASSCF)[52] method with second order perturbative energy corrections (CASPT2).[53] We examined energy transfer channels due to stretching of the XH bonds (X = N, O), which are known to play an important role in radiationless decay of heteroaromatic molecules.[54–57] We also investigated pathways involving out-of-plane distortions of the molecule. We compared these different relaxation pathways and interpreted the observed frequency-dependent deactivation of DHI[42] in light of our calculations.

3.2 Methods

Ground-state geometries of different conformations of DHI were optimized using the Møller–Plesset second-order perturbation (MP2)[22] method and the cc-pVDZ basis set.[58] For the calculation of vertical excitation energies and oscillator strengths, two methods were used: equation of motion coupled cluster singles and doubles (EOM-CCSD)[59, 60] and also the multistate (MS) version of CASPT2.[28][31] C_s symmetry was used. Since in a MS-CASPT2 calculation all the states must belong to the same irreducible representation of a point group, we used the following procedure to include the lowest seven singlet states (including the ground state), which belong to different irreducible representations of the C_s point group. First, using C_s symmetry, a state averaged (SA) CASSCF calculation was performed with three $A^{'}$ and four $A^{''}$ states corresponding to the lowest seven singlet states. Then, explicit C_s symmetry was dropped, but the zeroth-order wave function and orbitals for the MS-CASPT2 calculation were taken from this SA-CASSCF calculation. Because of the absence of explicit symmetry, all the seven states now trivially corresponded to the same irreducible representation. A level shift of 0.3 au was used for convergence of the MS-CASPT2 calculation. We used an active space consisting of 10 electrons in 10 orbitals (10,10) consisting of three pairs of $\pi - \pi^{*}$ and two pairs of $\sigma - \sigma^{*}$ orbitals (Figure 3.2). The geometry optimization of excited states S_1 and S_2, conical intersections (CIs) in the plane of the molecule, and corresponding pathways related to nonradiative decay of the first bright state were calculated using seven singlet states (three $A^{'}$ and four $A^{''}$ states) in the SA-CASSCF wave function with the explicit use of C_s symmetry. The same active space as that in the vertical energy calculation was used. Pathways for nonradiative decay of the first bright state were calculated with a combination of two approaches: linearly interpolated internal coordinates (LIIC) and relaxed scan calculations. In the former approach, geometries connecting critical points are obtained by interpolating them in internal coordinates at linear intervals, and energies of relevant electronic states are calculated at these geometries. In the latter approach, one internal coordinate (usually that which changes significantly between the critical points and called the driving coordinate) is held fixed at different intermediate values, and for each fixed value of the driving coordinate all other coordinates are obtained by optimizing the energy of the relevant electronic state. Dynamical correlation was included using the state specific (SS) CASPT2 method with a level shift of 0.3 au, using the SA-CASSCF reference wave function. Similar SA-CASSCF calculations with the lowest four states ($2A^{'} + 2A^{''}$) were also performed to compare the effect of state averaging. Non-planar CIs of the first bright state ($1^1\pi\pi^*$) with the ground state, and of the second bright state ($2^1\pi\pi^*$) with the first bright state ($1^1\pi\pi^*$) state, were optimized. The SA-CASSCF method was used to locate these geometries and perform corresponding LIIC calculations. In the case of the $1^1\pi\pi^*/\mathrm{S}_0$ CI, two pertinent states (S_0 and $1^1\pi\pi^*$) were used in the state-averaged wave function, and for the $2^1\pi\pi^*/1^1\pi\pi^*$ CI the two pertinent states in addition to the ground state (S_0, $1^1\pi\pi^*$, and $2^1\pi\pi^*$) were used. An active space for these calculations consisted of 10 electrons in 10 orbitals (10,10), which includes four pairs of $\pi - \pi^*$ orbitals and a pair of $\sigma - \sigma^*$ orbitals. As compared to the previous (10,10) active space, a $\pi - \pi^*$ pair is included in lieu of a $\sigma - \sigma^*$ pair for a correct description of non-planar structures, where the π-system is distorted. Dynamical correlation was included using the SS-CASPT2 method[29][30] (level shift 0.3 au) for the LIIC calculations. The EOM-CCSD, CASSCF, and CASPT2 calculations were performed using the 6-31++G** basis set.[61] The Molpro 2012 quantum chemistry software was used for all calculations.[34][35]

3.3 Vertical Excitation Energies

The optimized global minimum structure of the ground state of DHI is shown in Figure 2a, where the oxygen atom at position 6 is the hydrogen bond donor, and the oxygen

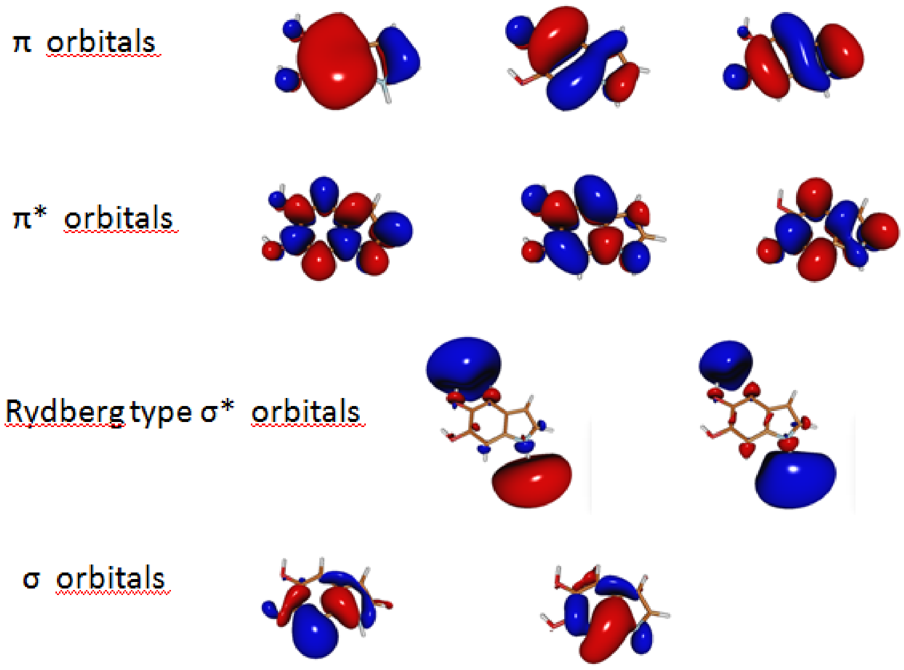


Figure 3.2: Active space containing 10 electrons in 10 orbitals

atom at position 5 is the hydrogen-bond acceptor. This is taken to be the geometry from which electronic excitation takes place upon photo-absorption. Addition of diffuse functions to the basis set does not make a significant difference in the calculated optimized structure.

Calculated vertical excitation energies and oscillator strengths at the ground-state geometry are presented in Table 3.1. On the one hand, from the oscillator strength values, the bright states can be identified to correspond primarily to electronic transitions from an occupied π to an unoccupied π^* orbital ($\pi\pi^*$ transitions). On the other hand, the dark states correspond primarily to $\pi\sigma^*$ transitions. We have investigated why the $\pi\pi^*$ states have a significant oscillator strength, while the $\pi\sigma^*$ states have negligible oscillator strengths.

Note that the σ^* orbitals associated with the lowest–lying dark states are Rydberg-type orbitals. For the correct description of these states it is necessary to use diffuse

Table 3.1: Calculated Vertical Excitation Energies (ΔE) and Oscillator Strengths (f) of DHI[a]

	MS-CASPT2/6-31++G**		EOM-CCSD/6-31++G**		CC2/aug-cc-pVTZ[b]	
Transition	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f
$1^1\pi\pi^*$	**4.61**	**0.0476**	**4.58**	**0.1161**	**4.39**	**0.141**
$1^1\pi\sigma^*$	5.01	0.0023	4.76	0.0017	4.67	0.0001
$2^1\pi\sigma^*$	5.19	0.0005	4.84	0.0032	4.77	0.0037
$3^1\pi\sigma^*$	5.58	0.0029	5.13	0.0003		
$4^1\pi\sigma^*$	5.86	0.0055	5.17	0.0002		
$2^1\pi\pi^*$	**6.37**	**0.2544**	**5.24**	**0.1045**	**4.83**	**0.067**

[a] The values for the bright states are shown in bold. [b] Values as reported in Ref [48].

Table 3.2: Effect of diffuse functions on the vertical excitation energies calculated using the EOM-CCSD method. Energy values are in eV.

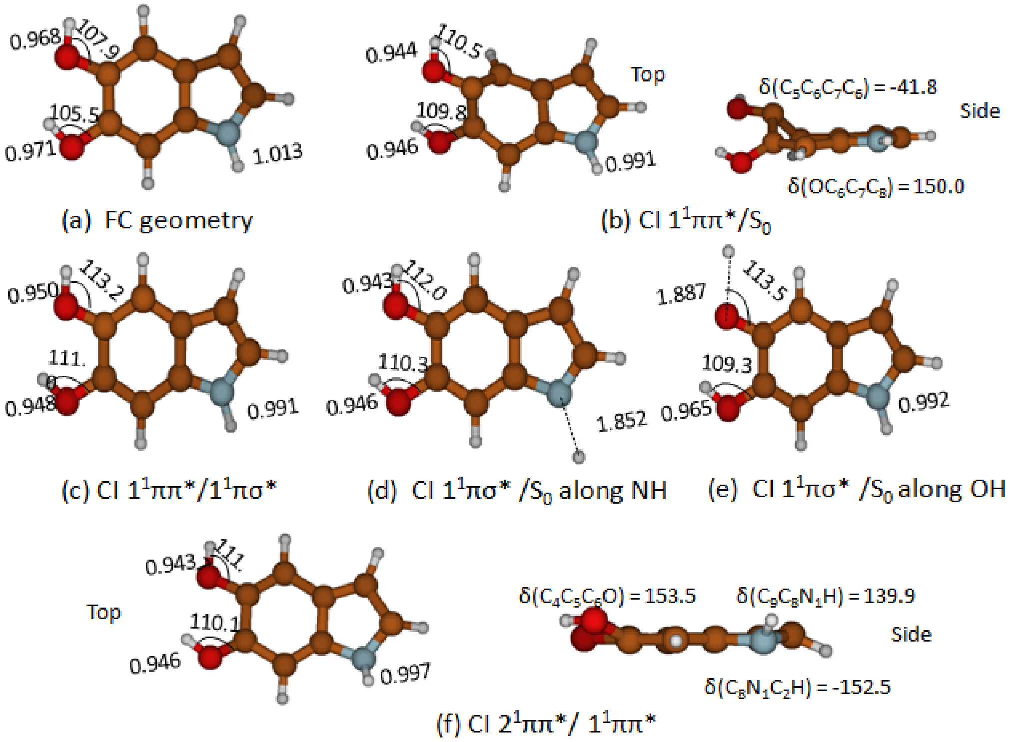
Transition	6-31++G**	6-31G*
$1^1\pi\pi^*$	4.58	4.74
$2^1\pi\pi^*$	5.24	5.51
$1^1\pi\sigma^*$	4.76	6.86
$2^1\pi\sigma^*$	4.84	7.34
$3^1\pi\sigma^*$	5.13	7.63
$4^1\pi\sigma^*$	5.17	7.89

[a] The values for the bright states are shown in bold.

Table 3.3: Calculated Vertical Excitation Energies (ΔE) and Oscillator Strengths (f) of DHI[a] using (10,12) active space

	MS-CASPT2/6-31++G**		EOM-CCSD/6-31++G**	
Transition	ΔE (eV)	f	ΔE (eV)	f
$1^1\pi\pi^*$	**4.54**	**0.0395**	**4.58**	**0.1161**
$1^1\pi\sigma^*$	4.89	0.0028	4.76	0.0017
$2^1\pi\sigma^*$	5.03	0.0011	4.84	0.0032
$3^1\pi\sigma^*$	5.31	0.0004	5.13	0.0003
$4^1\pi\sigma^*$	5.34	0.0035	5.17	0.0002
$2^1\pi\pi^*$	**5.57**	**0.0084**	**5.24**	**0.1045**

[a] The values for the bright states are shown in bold.



(a) FC geometry

(b) CI $1^1\pi\pi^*/S_0$

(c) CI $1^1\pi\pi^*/1^1\pi\sigma^*$

(d) CI $1^1\pi\sigma^*$ /S_0 along NH

(e) CI $1^1\pi\sigma^*$ /S_0 along OH

(f) CI $2^1\pi\pi^*$/ $1^1\pi\pi^*$

Figure 3.3: (a) Ground-state equilibrium geometry and conical intersection geometries between (b) $1^1\pi\pi^*$ state and ground state, (c) $1^1\pi\pi^*$ and $1^1\pi\sigma^*$ states, (d) $1^1\pi\sigma^*$ state and ground state along the NH stretch coordinate, (e) $1^1\pi\sigma^*$ state and ground state along the OH stretch coordinate and (f) the two lowest bright states. Bond lengths are in angstroms, and angles are in degrees.

functions in the basis set. In fact, if diffuse functions are not included, these Rydberg states do not appear among the lowest excited states. The energies of the $\pi\sigma^*$ states however were found to not change significantly whether diffuse functions were included or not (Table 3.2).

The vertical excitation energies calculated with MS-CASPT2 have the same ordering as the EOM-CCSD values, although the actual values are slightly greater. With more valence–type orbitals in the active space (five pairs of $\pi - \pi^*$ orbitals instead of three), the MS-CASPT2 vertical energies show better agreement with the EOM-CCSD values (Table 3.3) but require a relatively large level shift of 0.4 au for convergence. Moreover, including five pairs of $\pi - \pi^*$ orbitals along with other necessary σ-type orbitals in the active space is computationally impractical for a large number of calculations and was therefore not used for other geometries. The use of C_s symmetry for all vertical energy calculations was necessary to obtain consistent values across the MS-CASPT2 and EOM-CCSD approaches. Imposing symmetry helps in convergence to correct solutions by restricting the optimization space of the wave function, without compromising on the flexibility of the space required for the correct description of the wave function.

Interestingly, the Rydberg-type σ^* orbitals characterizing the lowest few excited states are located around the NH and 5-OH bonds, but not around the 6-OH bond (Figure 3.2). The 6-OH σ^* orbital is much higher in energy. This destabilization appears to be related to the participation of 6-OH in hydrogen-bond formation. To investigate this, we studied a phenol dimer model system, where we start with a hydrogen-bonded dimer and then move the monomers apart to see the effect of distance on the canonical Hartree–Fock orbital energies of the two OH σ^* orbitals (the 6-31++G** basis set was used). On the one hand, Figure 3.4 shows that at large O–H distances, the energies of the σ^* orbital of the donating and accepting hydroxyl groups are degenerate as expected. On the other hand, at shorter O–H distances, where hydrogen-bond interactions are present, the energy of the σ^* orbital on the donating OH bond (σ^*_d) increases, while

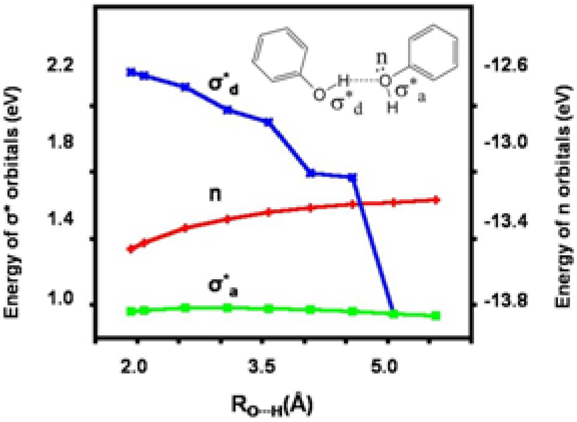


Figure 3.4: Energies of σ^* orbitals of donor and acceptor OH groups in phenol dimer along donor-acceptor distance. Also energy of lone pair on acceptor oxygen atom is plotted.

the energy of the other σ^* orbital (σ_a^*) does not change significantly. This trend can be rationalized based on molecular orbital interaction of the nonbonding n orbital on the acceptor oxygen atom with the empty σ^* orbital along the donor OH bond, giving rise to a stabilized bonding (contributing to formation of the hydrogen bond and having primarily acceptor oxygen n character) and a destabilized anti–bonding (having primarily σ^* character). Figure 3.4 shows the energy stabilization of the n orbital accompanying the destabilization of the σ_d^* orbital. Consequently, the σ_d^* orbital with a relatively high energy is not associated with the lower $\pi\sigma^*$ transitions of the phenol dimer and by extension with the lower energy transitions of DHI.

3.4 Excited-State Equilibrium Geometries and Conical Intersections

Geometry optimization for the lowest $\pi\pi^*$ and lowest $\pi\sigma^*$ states (S_1 and S_2 states, respectively, at the ground-state geometry) resulted in geometries similar to the ground-state equilibrium structure. The energy of the $1^1\pi\pi^*$ minimum calculated at the CASPT2 level is 4.25 eV with respect to the equilibrium ground-state energy. This can be compared to the observed emission maximum of DHI, which is at 3.2 eV.[42] The difference between the two values can be attributed at least partially to the solvatochromic shift in the experimental results, which are performed in water (calculated dipole moments of the ground and $1^1\pi\pi^*$ are 2.5 and 3.0 D, respectively) and also to the lack of inclusion of zero-point effects in the calculation. In the case of the $\pi\sigma^*$ state, finding a minimum-energy structure is in contrast to previous single reference studies, which suggested that the $\pi\sigma^*$ state has a barrierless dissociative potential energy surface for detachment of the hydrogen atom from the OH group at position 5.[48]

A search for the minimum-energy CI geometry between the first bright state ($1^1\pi\pi^*$) and the energetically close by $1^1\pi\sigma^*$ state was performed starting from the ground-state geometry. This resulted in a CI structure shown in Figure 3.3c. The main difference with respect to the FC geometry is an increase in the COH angles. Orbital analysis at this CI geometry showed that the σ^* orbital in the $\pi\sigma^*$ state continues to be a combination of σ^* orbitals along the NH and OH bonds, like it was at the FC geometry. Given that the antibonding σ^* orbital has a node along the center of the XH bonds, one can expect stabilization of $\pi\sigma^*$ state due to stretching of these bonds. At the same time, these geometry perturbations would destabilize the ground state and likely lead to an intersection between the two states. Hence, we performed CI searches between the lowest $\pi\sigma^*$ state and the ground state by starting from two structures with the NH and OH bonds stretched separately. Figure 3.3d,e shows the CI geometries obtained, where the NH and OH bonds are significantly stretched. These CIs can act as funnels for radiationless energy transfer of the molecule from the excited $\pi\sigma^*$ state.

A CI search between the $1^1\pi\pi^*$ and ground state was initiated by starting with several distorted structures by twisting the C_4–C_5 and C_6–C_7 bonds of DHI, similar to the procedure followed to obtain the out-of-plane CIs in adenine.[62] These ring-puckered initial distortions led to three CIs. The lowest energy CI (Figure 3.3b) has an energy that is 0.46 eV higher than the $1^1\pi\pi^*$ state at the FC geometry, while the others have higher energies. The C_5 moves significantly out of the plane, and the dihedral angle $\delta(C_5C_6C_7C_8) = 41.8^o$. Unsubstituted indole has a similar non-planar CI, with the C_6 going out of the plane,[63] and the dihedral angle $\delta(C_9C_8C_7C_6) = 32.9^o$.[64]

A CI between the two $\pi\pi^*$ bright states was also found, and this involves not only out-of-plane motions of hydrogen atoms attached to the nitrogen atom and the neighboring carbon atom but also the hydroxyl group attached to carbon C_6 (Figure 3.3f). The CI is accessible with small geometry changes with respect to the FC geometry.

3.5 Pathways for Nonradiative Decay

In this subsection, we present the pathways for nonradiative decay after photoexcitation to the first two bright states — S_1 ($1^1\pi\pi^*$) and S_6 ($2^1\pi\pi^*$). For the $1^1\pi\pi^*$ state, there is a CI with the S_2 ($1^1\pi\sigma^*$) state near the FC geometry. Since there is no single dominant coordinate responsible for this CI to be accessed after vertical excitation, we performed LIIC calculations to find the path connecting the FC geometry to the CI. As in the case of several heteroaromatic molecules,[55][56][65] this is expected to be the predominant path along which the molecular population transfers from the bright $\pi\pi^*$ to the dark $\pi\sigma^*$ state. Another LIIC calculation was performed to connect this CI to the minimum on the $\pi\sigma^*$ state. From this minimum-energy structure, stretching of the NH and OH bonds leads to two other CIs. We performed relaxed scans on the σ^* state for finding pathways leading to these CIs with the ground state.

The SA-CASSCF energies of the lowest seven states along the two pathways for nonradiative decay after photoexcitation to the $1^1\pi\pi^*$ state are shown in Figure 3.5. Since the lowest three states are particularly significant for understanding the de-excitation dynamics of the first bright state, the energies of these states along the deactivation

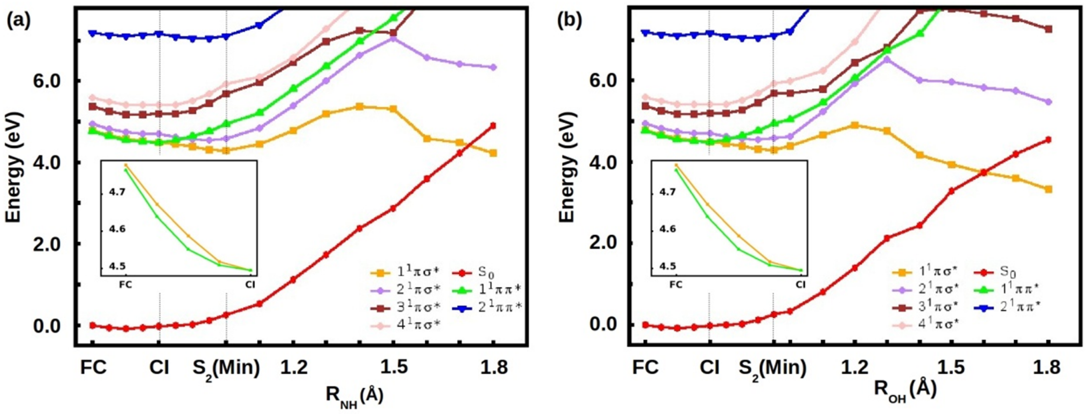


Figure 3.5: Pathways for nonradiative decay of first bright state calculated using the 7 SA-CASSCF method. In both (a) and (b), the first section of the abscissa represents the LIIC path connecting the FC and S_1/S_2 CI geometries, and the next section represents the LIIC path connecting the CI with the minimum energy structure on the $\pi\sigma^*$ state. The final section corresponds to a relaxed scan on the $\pi\sigma^*$ state connecting its minimum energy geometry to its conical intersection with the S_0 state along the (a) NH and (b) OH stretching coordinates. Insets in both (a) and (b) display zoomed part of LIIC path connecting FC and S_1/S_2 CI, with energies for S_1 and S_2 plotted.

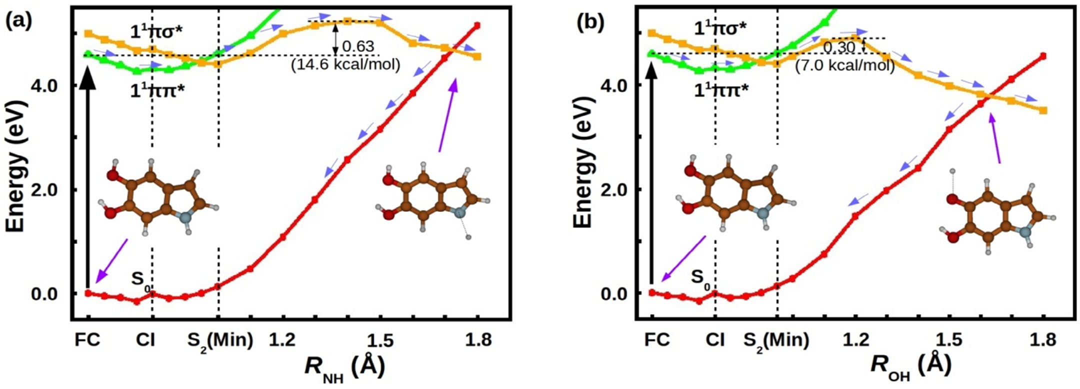


Figure 3.6: Potential energy profiles calculated using CASPT2 method along CASSCF geometries. In both (a) and (b), the first section of the abscissa represents the LIIC path connecting the FC and S_1/S_2 CI geometries, and the next section represents the LIIC path connecting the CI with the minimum energy structure on the $\pi\sigma^*$ state. The final section corresponds to a relaxed scan on the $\pi\sigma^*$ state connecting its minimum energy geometry to its conical intersection with the S_0 state along the (a) NH and (b) OH stretching coordinates.

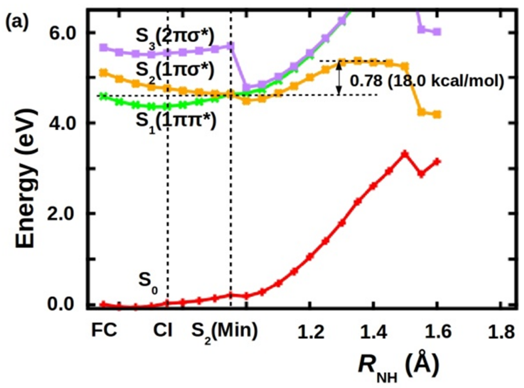


Figure 3.7: Potential energy profiles calculated using CASPT2 method along 4 SA-CASSCF geometries. In both (a) and (b), the first section of the abscissa represents the LIIC path connecting the FC and S_1/S_2 CI geometries, and the next section represents the LIIC path connecting the CI with the minimum energy structure on the $\pi\sigma^*$ state. The final section corresponds to a relaxed scan on the $\pi\sigma^*$ state connecting its minimum energy geometry to its conical intersection with the S_0 state along the (a) NH and (b) OH stretching coordinates.

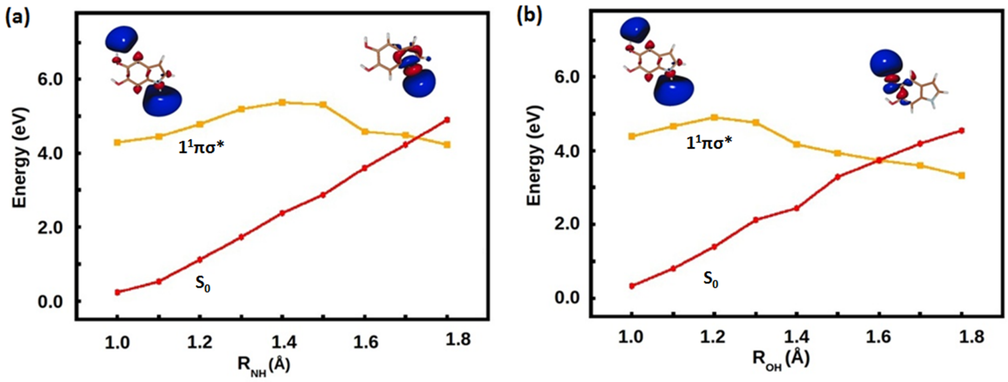


Figure 3.8: (a) and (b) show the change in the nature of lowest energy σ^* CASSCF orbital (from Rydberg to valence) along NH and OH bond stretching for relaxed scan along $\pi\sigma^*$ state, respectively. The energies of $\pi\sigma^*$ and the ground states are plotted along this path calculated with 7SA-CASSCF method with (10,10) active space.

pathways are calculated including dynamical electron correlation: the CASPT2 energies for these states are presented in Figure 3.6. Note that the S_1/S_2 intersection geometry at the CASPT2 level changes slightly from the abscissa value calculated at the SA-CASSCF level and marked as CI in the figure. Similar potential energy curves calculated with SA-CASSCF with the lowest four states followed by CASPT2 corrections are shown in Figure 3.7.

The $\pi\sigma^*$ state has a minimum-energy geometry and an energy barrier along the NH and OH paths to access the CI with the ground state. This behavior is typical of $\pi\sigma^*$ states in heteroaromatic molecules such as pyrrole, indole, and phenol, and it has been widely discussed in the literature.[54][55][56][65] The nature of the σ^* is of 3s Rydberg type in the FC region and changes to valence σ^* as the OH or NH bond is stretched and eventually to a 1s orbital on the hydrogen atom (Figure 3.8). The change of state character is reflected in the double well shape of the potential energy surface. The effective energy barriers along the NH and OH bond-stretching paths, that is, the difference between the energy after vertical excitation to the $1^1\pi\pi^*$ state and the maxima on the $1^1\pi\sigma^*$ state along the two paths, are 0.63 and 0.30 eV, respectively. These values are similar (0.78 and 0.36 eV) for a state-averaged calculation with the lowest four states. The trend in these barrier heights — NH greater than OH — is the same as in 5-hydroxyindole, where the barrier at the EOM–CCSD level with respect to the lowest point on the $1^1\pi\sigma^*$ curve along the OH path is just in the excess of 0.1 eV, while that along the NH path is 0.46 eV.[66] Also in 5-hydroxyindole, experiments led the authors to speculate that there is no transfer of hydrogen from hydroxyl group at position 5 to the neighboring carbon, which is consistent with the pathways obtained by us for DHI.[66]

Accessing the $1^1\pi\sigma^*/S_0$ CI after excitation to the first bright state involves significant stretching of the NH and OH bonds. To explore the importance of coupling of other vibrational modes to these stretches, we performed rigid scans of the energies of the

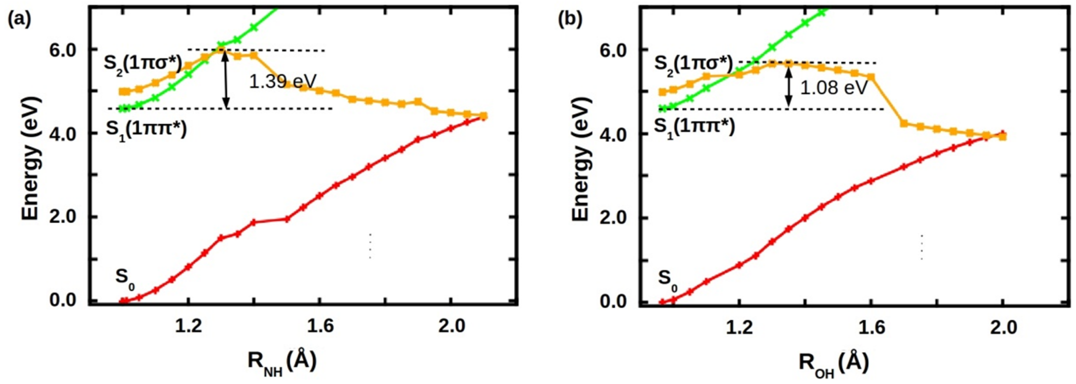


Figure 3.9: Rigid scan of potential energy curves along (a) NH and (b) OH stretching coordinates with CASPT2/7SA-CASSCF method. The lowest three states are plotted.

lowest seven states by stretching only the NH and OH bonds starting from the FC geometry and keeping other internal coordinates fixed. The same level of theory as for the relaxed scans was used (Figure 3.9). The barrier with the rigid scan is higher than the relaxed scan. The major distortions accompanying the NH and OH stretches are in the C5OH and C6OH angles. For example, at the constrained bond length of NH = 1.3 Å, the rigid scan has $\theta(C_5OH) = 107.9^o$ and $\theta(C_6OH) = 105.5^o$, which are just the values of these angles at the FC geometry. These change on relaxation to 112.7^o and 111.4^o, respectively. Similarly, at the constrained bond length of OH = 1.3 Å, $\theta(C_5OH)$ and $\theta(C_6OH)$ change on relaxation from the FC values to 120.2^o and 108.8^o, respectively. These distortions are reflective of the σ^* orbitals having a mixed character with NH and 5OH components, and furthermore the presence of hydrogen bonding between the 5OH and 6OH.

The CASPT2 energies of the $1^1\pi\pi^*$ and ground state along the LIIC path connecting the FC geometry with the $1^1\pi\pi^*/S_0$ CI was calculated (Figure 3.10). The energy of the CI is 0.46 eV with respect to the $1^1\pi\pi^*$ vertical excitation energy, and there is no further barrier on the path to access this intersection.

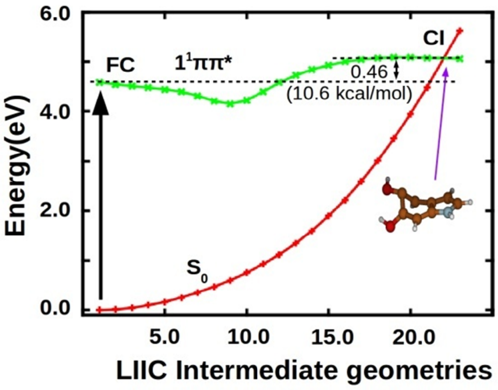


Figure 3.10: LIIC pathway connecting FC geometry to ring-puckered $1^1\pi\pi^*/S_0$. CASPT2 single point energy calculations were performed.

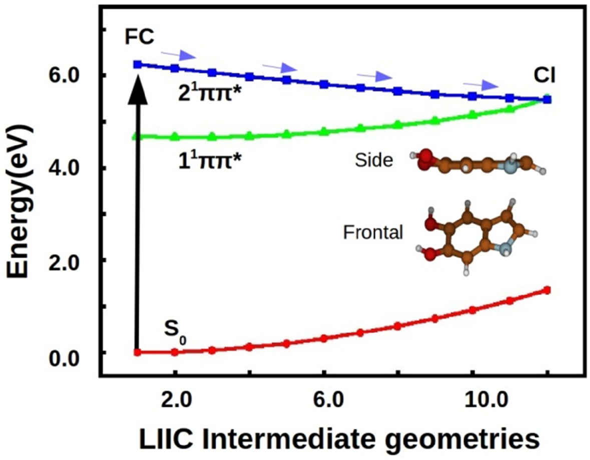


Figure 3.11: LIIC pathway con3necting FC geometry to CI between two bright states. CASPT2 single point energy calculations were performed.

For understanding the nonradiative decay mechanism of the second bright state (S_6), we calculated CASPT2 energies along the path connecting the FC geometry to the CI with the first bright state (Figure 3.11). The pathway to the non-planar CI is energetically downhill and barrierless. This suggests that on photoexcitation of the molecule to the second bright state, internal conversion to the first bright state takes place at an ultrafast time scale.

On the basis of the calculations of pathways, we are in a position to propose deactivation mechanisms of DHI starting from the S_1 ($\pi\pi^*$) state, which could be reached either by direct photoexcitation or by internal conversion from the second bright state. From the state, DHI accesses the conical intersection with the nearby $\pi\sigma^*$ state, thereby transferring molecular population to this $\pi\sigma^*$ state. Then the molecule can move toward the $\pi\sigma^*$ minimum and potentially access the CI with the ground state by stretching the OH or NH bonds based on the available vibrational energy in the respective modes. Comparing the vertical energy of S_1 with the energy of the highest energy point on the $\pi\sigma^*$ pathway suggests that the excitation to higher vibrational states on the S_1 state is required for the molecule to have sufficient energy to overcome the barrier for accessing the conical intersection with the ground state. Similarly, accessing the out-of-pane $1^1\pi\pi^*/S_0$ CI, which is higher in energy than the vertical energy of S_1, requires excitation to higher vibrational states.

The present study assists in interpreting the time-resolved fluorescence experiments of Huijser, Pezzella, and Sundström.[42] In these experiments, a marked excitation energy dependence of the fluorescence lifetime is observed. The lifetime decreases when the excitation frequency is increased. The long lifetime on excitation to the first bright can be rationalized based on the barriers present for accessing the $1^1\pi\sigma^*/S_0$ CI and also the $1^1\pi\pi^*/S_0$ CI. Higher-frequency excitation, besides producing a larger fraction of DHI in the second bright state, can excite the molecule to higher vibrational states of the first bright state and hence provide energy to cross the barrier and lead to a shorter excited-state lifetime. For the second bright state, there is a possible competition and consequent branching between two decay channels, both of which are likely to occur on the subpicosecond time scale: one, a barrierless internal conversion to the lower $1^1\pi\pi^*$ state through a CI, and two, an electron transfer to the solvent to form the DHI radical cation. The latter, that is, the radical cation formation, was attributed with the short lifetime after excitation to the $2^1\pi\pi^*$ state by Sundström and co-workers but is beyond the scope of the present study.

3.6 Conclusions

The study of the deactivation mechanisms of DHI using multi-reference methods provides insights that are significantly different from earlier studies with single-reference methods.[48] After excitation to the lowest optically bright state ($1^1\pi\pi^*$), the molecule undergoes internal conversion to the $1^1\pi\sigma^*$ state through an easily accessible CI. The $1^1\pi\sigma^*$ state is found to have a minimum-energy geometry, in contrast to earlier suggestions of barrierless hydrogen migration from the hydroxyl group to the neighboring carbon atom occurring on this state. Barriers were found for nonradiative decay of the $1^1\pi\sigma^*$ state along NH and OH stretching coordinates to relax to the ground state. The $1^1\pi\pi^*$ state additionally has a ring puckered CI with the ground state, but it is higher in energy than the vertical excitation energy of $1^1\pi\pi^*$ and effectively has a barrier for access. Excitations to higher vibrational states can overcome these barriers and explain the experimentally observed excitation energy dependence of the excited-state lifetimes.[42]

For the second optically bright state ($2^1\pi\pi^*$), internal conversion to the $1^1\pi\pi^*$ state is possible through a barrierless CI. However, there is also the possibility of electron transfer to water,[42] about which no conclusions can be drawn from the present investigation, which is on isolated DHI. Further studies on the DHI radical cation and the interaction of DHI with solvent molecules are needed.

A higher barrier was observed for accessing the $\pi\sigma^*/\mathrm{S}_0$ CI along the NH bond stretching as compared to OH bond stretching. This may suggest that energy transfer may take place predominantly through the OH pathway, but effects such as gradients and initial energy available in particular modes need to be taken into account to make accurate predictions. At the $\pi\sigma^*/\mathrm{S}_0$ CI, the molecule on the $\pi\sigma^*$ state can either nonradiatively decay to the ground state or continue on the repulsive part of the potential energy surface and dissociate to form DHI radical and a free hydrogen atom. These processes would depend on the local topology of the potential energy surface, wavepacket velocity, and the coupling between states. Nonadiabatic dynamical studies are necessary to understand the branching ratios of these competing processes.

In DHI, the antibonding σ^* orbital corresponding to the hydrogen-bonded OH bond (6-OH) was found to be much higher in energy than the 5-OH σ^* orbital. This can be understood based on molecular orbital interaction of the σ^* with the nonbonding orbital on the hydrogen-bond acceptor. Further investigations are required to establish whether this is a general feature of hydrogen-bonded systems. The relatively long excited-state lifetime of DHI proposed from our study, along with previous experimental evidence, suggests that the photoproperties of isolated monomeric DHI cannot explain the picosecond deactivation properties of eumelanin. Dimers and polymers of DHI, and their interactions with other building blocks of eumelanin, need to be studied to understand such properties.

Chapter 4

Solvation-Driven Keto-to-Enol Tautomerization of Barbituric Acid

4.1 Introduction

Barbituric acid (BA) is a proposed proto-RNA nucleobase.[67–69] Proto-RNA has been speculated to be a precursor of RNA and contains nucleobases which form nucleosides more easily than the current nucleobases. BA is a pyrimidine-based molecule structurally similar to the RNA nucleobases uracil and cytosine. Unlike these contemporary nucleobases, BA shows the formation of glycosidic bonds to produce nucleosides and nucleotides via simple chemical routes.[70, 71] It has also been shown re-cently that BA when mixed with melamine in water readily forms a base pair, analogous to the Watson-Crick base pairs, which is a property used for information transfer in living systems.[70] These observations are useful for understanding the evolution of RNA on prebiotic earth, with water as the abundant solvent. Besides playing a possible role in the evolution of life, BA is well known as the parent molecule of barbiturate drugs, which act on the central nervous system and have a wide variety of pharmacological applications.

BA can undergo tautomerization due to hydrogen transfer from methylene or amine groups (see Figure 4.11a) and shows as many as 16 tautomers.[72] For nucleobases, tautomerization plays an important role when present in RNA or DNA. In fact, the structure and function of RNA and DNA is crucially determined by the dominant tautomeric forms of their constituent nucleobases. This was recognized as early as 1953 by Watson and Crick, who based their proposed structure of DNA on the assumption that the most plausible tautomeric form of the four nucloebases is keto.[73] Minor tautomeric forms have been suggested to lead to mutation of DNA during replication[74–77] and transcription.[78] In the case of RNA, minor tautomeric forms of nucleobases are proposed to play an essential role in the catalytic role of ribozymes and also in the recognition of ligands by riboswitches.[79–82] The tri-keto form of isolated BA, henceforth referred to as the keto form or keto tautomer, is the most stable tau-tomer (Figure 4.1a) and the enol form the second most-stable tatutomer (Figure 4.1bb).[72, 83–86] The keto form is 11 kcal/mol more stable than the enol form based on these studies. On the other hand, in the solid state, the enol form of BA has been found to be more stable than the keto form.[87, 88] BA is found to spontaneously convert from the keto to the enol form. The greater stability of the enol form has been attributed to the formation of addi-tional inter-molecular hydrogen bonds as compared to the keto form in the crystal leading to a more stable lattice energy.[87, 88] In aqueous solution, which is the most biologically relevant environment for BA, there is no definitive understanding about its most stable tautomer. Studies modeling the aqueous environment as a dielectric continuum suggest that the isolated enol tautomer gets more stabilized in water than the isolated keto tautomer by 3 kcal/mol, but the keto form continues to be the most stable.[84, 89] In a similar study, performed with a different solvation model, it was found that solvation does not affect the relative stability of keto and enol tautomers.[83] Effect explicit water molecule on the activation barrier energy studied semi-empirical methods has suggested

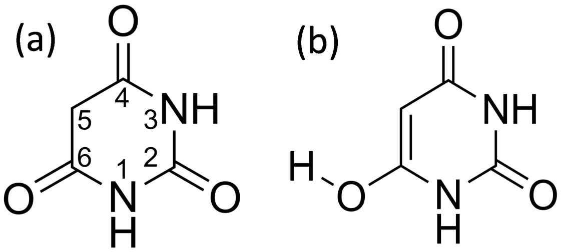


Figure 4.1: Relevant tautomers of barbituric acid: (a) keto and (b) enol.

reduction in barrier height by 26.5 kcal/mol.[82] In a similar study performed using quantum chemical method, it was found that water molecule taking part in hydrogen transfer at the transition state reduced keto-to-enol transition barrier height by 32.7 kcal/mol, which resulted in keto-to-enol transition barrier height of 30.7 kcal/mol.[72] Recently, the effect of explicit water molecules on the relative stabilities of BA tautomers has been studied using DFT calculations by considering 3 water molecules complexed with BA.[**ba˙2106**] Since BA has the possibility of forming several hydrogen bonds with water, considering only 3 solvating water molecules are not enough to draw conclusions about the relative stability of its tautomers in solution.

There are several studies on the effect of hydration on the stability of different tautomeric forms of nucleobases other than BA. To study the relative stability between keto and enol tautomers in uracil and 5-bromouracil, Monte-Carlo calculations have been performed with 400 water molecules, to obtain one suitable input structure for quantum optimization of solute with closest 50 water molecules.[89] Hobza and co-workers in their studies on relative stabilities of tautomers of all DNA and RNA nucleobases in the aqueous phase, performed MD simulations of nucleobase along with 2 water molecules to generate structures for classical optimizations, followed by clustering.[90–93] The study showed that some tautomers which are not the most stable in the gas phase can get stabilized due to interaction with water molecules. The studies Danilov and Hobza account for explicit solvation effects. However, in the former study only one structure of cluster containing solute and 50 water molecules is considered, which does not account for enough sampling. On the other hand clusters examined in the later study contain just 2 water molecules, which does not adequately describe hydration effects. Genetic Algorithms have also been applied for ideintifying golbal minimum of molecular clusters.[**gas1**, 94, 95] Such algorithms could also be applied for studying the relative stability of clusters containing BA-tautomer and interacting water molecules.

Various experimental methods have been employed to study tautomerization in DNA/RNA nucleobases. Fluorescence spectroscopy of thymine tautomers have shown possible existence of the less stable enol form in the aqueous phase.[96] Similarly, for 5-chloro-uracil, using fluorescence spectroscopy it was shown that enol tautomer contributes to the spectra at pH4.[97] FTIR experiments on cytosine-water complexes in Argon-matrix have shown stabilization of iso-cytosine tautomer.[98]

In this chapter, we present a detailed study of the change in the relative stability of keto and enol tautomers of BA with increasing extent of hydration. A computational procedure based on a combination of MD simulations and quantum chemical calculations is used, which treats water explicitly, accounts for statistical effects and is computationally practical. The results suggests that the enol tautomer of BA gets significantly stabilized due to interaction with water, and is in fact the more stable tautomer of BA in the aqueous phase as opposed to isolated BA where the keto form is more stable. IR and NMR experiments independently support this conclusion.

4.2 Computational Details

DFT was applied for optimizing the minimum energy geometries of isolated BA-tautomers as well as clusters containing BA-tautomer and water molecules. The Becke 3-parameter Lee-Yang-Parr (B3LYP) hybrid functional was employed.[99, 100] The 6-31++G** basis set was used for the description of hydrogen bond interactions in solute-solvent clusters.

The Gaussian 09 package was used for all quantum chemical calculations.[101]

Molecular dynamics simulations for keto and enol tauto-mers of BA were performed using GROMOS 53A6 force field.[102] The force field parameters were obtained using the automatic topology builder (ATB).[103] The partial charges that are used to run MD trajectories, as obtained with ATB, are shown in Figure 4.2. Two sets of simulations were performed: one for keto tautomer and the other for enol tautomer. Same protocol was followed for both of these sets of simulations. A cubic box of length 21 Åwas created and the BA-tautomer at the center of the box. Energy was minimized using steepest descent method for 10000 steps.[104] The box was solvated with 323 SPCE water molecules[105] for enol and 322 SPCE water molecules[105] for keto. Energy of each system was minimized with steepest descent method for 10000 steps.[104] Then NVT simulation was performed for 10 ns at constant temperature 300 K using Nosé-Hoover thermostat[106, 107] with a coupling constant of 0.4 ps. Later, NPT simulation was performed for 30 ns at constant temperature 300 K using Nosé-Hoover thermostat[106, 107] and constant pressure 1 bar using Parrinello-Rahman barostat[108] with a coupling constant of 0.4 ps each. After NPT run box dimensions changed to 21.4 Åfor system with keto tautomer and 21.6 Åfor system with enol tautomer. Finally, we performed production run for 100 ns, using NVT conditions (as mentioned above). The time step in each simulation was 1 fs. Periodic boundary conditions were employed in all directions to mimic the bulk effect. Particle Mesh Ewald method[109] with a cut-off of 10 Åwas used for calculating electrostatic interactions. 10 Åwas applied for calculating van der Waals interactions. Molecular dynamics (MD) simulations were performed with the Gromacs 4.5 package.[110] Data obtained from production run was considered for the further analysis.

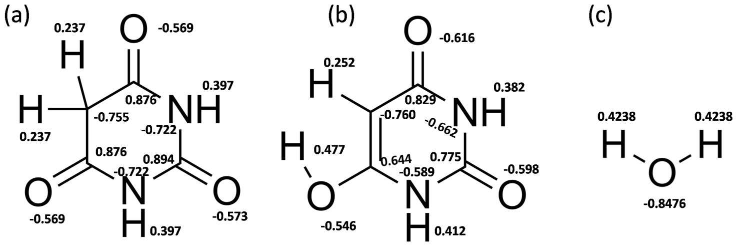


Figure 4.2: Partial charges used to run MD trajectories: (a) keto and (b) enol (c) water.

4.3 Results and Discussion

Geometry optimization of the two lowest energy tautomers of BA, namely the keto and enol tautomers, at the B3LYP/6-31++G**level of theory suggests that the keto form is more stable than enol by 9.91 kcal/mol. This is consistent with previous calculations have predict this relative stabilization to be 13.68 kcal/mol (B3LYP/6-31G*)[84] and 10.82 kcal/mol (B3LYP/D95**).[85] We also optimized the isolated molecules at the MP2/cc-pVDZ level of theory and found a similar relative stability of 10.52 kcal/mol. To account for the bulk effects of water on the stability of the tautomers the polarizable continuum solvation model (PCM) has been used. The relative stability of keto was to be 8.11 kcal/mol. This relative energy difference matches very well with 8.09 kcal/mol calculated at the B3LYP/IEF-PCM/6-311++G** level of theory.[111] As the effect of PCM corrections to relative stability is negligible we did not use PCM in further calculations.

4.3.1 Clusters with 1-4 Solvating Water Molecules

Starting geometries for the optimizations were chosen intuitively by adding water molecules which are hydrogen bonded to BA. Key structural difference between keto and enol tautomers is at the C_5 and C_6 positions. Thus, it is essential to see the effect of hydration by adding explicit water molecule(s) in C_5-C_6 and C_6-N_1 moieties. We performed geometry optimizations of clusters including BA molecule and water molecule(s) hydrogen bonded

BA at the C_6 position as presented in Figure 4.3. These clusters were obtained by manual optimizations of various initial geometries, which were chosen in such a way that the water molecule(s) is(are) interacting with the BA molecule via hydrogen bonding with carbonyl or hydroxyl groups at the C_6 position.

Addition of one water molecule is simple in the case of keto form since carbonyl group can only act as acceptor and water molecule can be situated near the C_5-C_6 or C_6-N_1 moiety. On the other hand for enol, the water molecule can act as hydrogen bond donor with the C_6-N_1 moiety or hydrogen bond acceptor with the C_5-C_6 moiety when interacting with the hydroxyl group. These optimized geometries are shown in Figure 4.3a. The upper panel of Figure 4.3a represents clusters of type 1 and lower panel represents clusters of type 2. We find that the hydroxyl group of enol tautomer does not act as hydrogen bond acceptor when interacting with one water molecule. Several structures were optimized with an initial structure having water molecule placed in an N_1-C_6 moiety of enol and as a hydrogen bond donor to the hydroxyl group but, all of these structures led to breaking of water-hydroxyl hydrogen bond and formation of hydrogen bond between N_1 and O of a water molecule.

Adding second water molecule to both the clusters of a BA-1 water molecule (Figure 4.3a) is possible in such a way that it can either form a hydrogen bond with the first water molecule or it forms a direct hydrogen bond with the carbonyl (keto) or hydroxyl (enol) groups. With this, there are three possibilities to add a second water molecule as shown in Figure 4.3b. Starting with three types of clusters of BA-2 water molecules (Figure 4.3b), the addition of a third water molecule leads to two possible types of clusters. These clusters are shown in Figure 4.3c. Finally, the addition of a fourth water molecule to the clusters of BA-3 water molecules leads to only one unique possible cluster as shown in Figure 4.3d. Table 4.1 presents average energies and dipole moments of clusters containing 1-4 water molecules with keto and enol tautomers. Clusters with enol tautomer are stabilized more than the clusters with keto tautomer on the addition

Table 4.1: Energies of clusters with 1–4 water molecules. Energies are given in $kcal/mol$ and dipole moments (DM) in Debye. Average energies and DMs for each cluster are presented. Zero is adjusted with respect to the lowest energy cluster for 1–4 water molecules.

Number of water	Number of types	Keto Energy[a]	Keto DM[a]	Enol Energy[a]	Enol DM[a]	Relative Stability
0	1	0.00	1.19	9.91	4.60	9.91
1	2	1.44	1.86	9.87	7.49	8.44
2	3	3.55	2.36	12.09	6.33	8.55
3	2	0.72	3.39	7.76	6.53	7.04
4	1	0.00	3.26	5.80	4.38	5.80

[a] Clusters have occurrence percentage higher than 50% (except for the case of cluster type for enol 511111 and 611211), which means that any other type of cluster in these cases is not as probable as the ones presented here. For enol 5-hydrogen bonds with water, the second largest abundant cluster is 510121 with percentage 20.73 and for enol 6-hydrogen bonds with water, second largest cluster is 611121 with percentage 21.59.

of water molecules. This additional stabilization of enol clusters is probably due to stronger hydrogen bonding of enol tautomer than keto tautomer with surrounding water molecules in clusters. It is necessary to add more water molecules to predict the effect of solvation on the relative stability of keto and enol tautomers in water. For adding 5^{th} and beyond water molecules to the system, based on intuition, there are many possible ways. As for example, a 5^{th} water molecule can be added in 4 different ways for keto, i.e. between N_1-C_2, C_2-N_3, N_3-C_4, and C_4-C_5 moieties. Also, generating clusters by adding 5^{th} water molecule in all possible ways would lead to high variations in energies, which in turn adds to the error in values for relative stability. Hence, initial geometries for quantum optimizations of clusters containing more than 4 water molecules cannot be chosen in a systematic way just based on intuition.

4.3.2 Solvation Effects with MD Simulation

MD simulations were performed to study the effect of solvation on the keto and enol tautomers. These simulations are useful for mimicking the bulk water environment

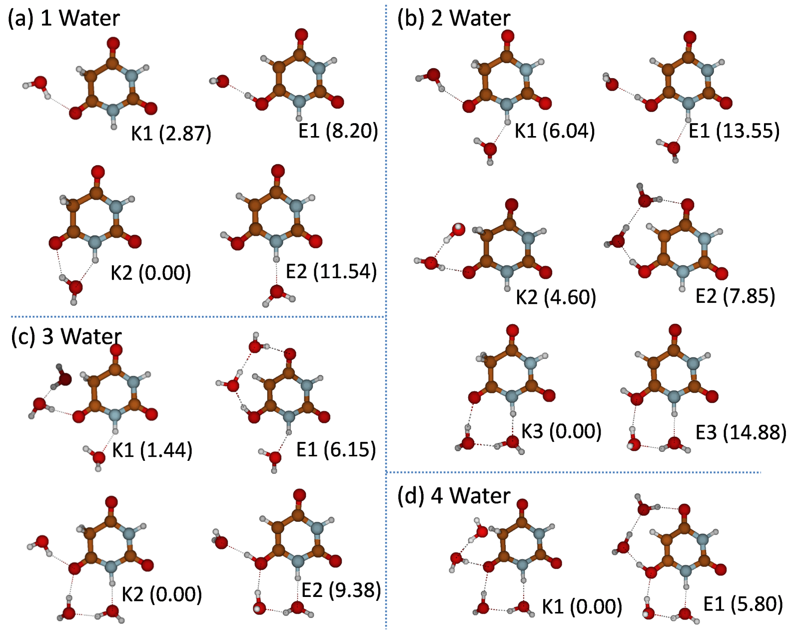


Figure 4.3: Clusters of BA with water molecules. Inclusion of 1, 2, 3 and 4 water molecules in (a), (b), (c) and (d) respectively. Structures on left are keto and right is enol in each group. Numbers near each cluster represent cluster type with num-ber in parenthesis showing relative energy (in kcal/mol) of the cluster with respect to lowest energy cluster in each group.

around the solute molecules.

In principle, relative stabilities of keto and enol tautomers can be calculated using MD. For this, we calculated the average potential energy of solute and 25 surrounding water molecule classically from MD simulations. The water molecules with center of mass closest to the solute were considered. However, as MD does not describe electronic rearrangement in the molecules, hence we cannot calculate internal energy classically. Secondly, force field description for keto and enol differs due to the presence of a hydroxyl group in enol, it adds to 1-4 interactions and dihedral energies. The difference in energy due these factors is very small (in orders of magnitude) compared with average potential energies of systems and hence errors associated with relative energy calculations are higher than relative energy calculated quantum chemically. For these reasons, calculating relative energy from classical MD simulation cannot give accurate results.

Choosing clusters containing solute and surrounded water molecules for performing quantum optimization is an arduous task. We performed single-point energy calculations for clusters chosen from MD trajectory. These clusters containing solute and surrounding water molecules were selected at regular intervals from the MD trajectory. Quantum chemical calculations of energy are highly sensitive to the orientations molecules in clusters. Thus, quantum chemical energies calculated in this fashion gave higher variations in relative energy as much as 10 kcal/mol.

For an appropriate comparison of relative stability, it is necessary that clusters that are chosen for quantum chemical calculations should be taken from their most probable occurrence. In MD simulations, trajectory spends higher time in orientations with higher free energy. Choosing such configurations as an input for further quantum chemical calculations would yield statistically meaningful results. A most common way of obtaining such most probable structure in MD is performing cluster analysis based on mean square displacement (MSD). However, as water molecules are highly mobile it is difficult to perform cluster analysis when water is the solvent. To remedy this, we propose a protocol for systematically choosing clusters with the higher number of water molecules. We did clustering based on the distribution of hydrogen bonds between solute and solvent molecules.

Figure 4.4 describes the normalized distribution plot of solute-solvent hydrogen bonds. Geometrical criteria to determine if hydrogen bond exists is same as than used in Gromacs software package.[110] This criteria is defined as following, a hydrogen bond is assumed to exist when donor–acceptor distance ≤ 3.5 Åand hydrogen–donor–acceptor angle is $\leq 30^o$. This distribution is obtained by first plotting histograms for solute-solvent hydrogen bonds for keto and enol tautomers separately and then normalizing the sum of the probability of dydrogen bond to unity. An average number of hydrogen bonds distribution between solute and solvent molecules is at 6.9-hydrogen bonds for enol and at 5.9-for keto. This distribution implies that one extra hydrogen bond is formed by the enol tautomer with water molecules than that by the keto tautomer. A number of solute-solvent hydrogen bonds having a relative population greater than 0.2 is considered to be most-probable event in the MD trajectory. Frames having 5, 6, and 7 solute-solvent hydrogen bonds for keto, and6, 7, and 8 solute-solvent hydrogen bonds for enol are most-probable clusters. For comparison of relative stability of keto and enol, we considered clusters having 5-to-8 solute-solvent hydrogen bonds. Further analysis of these most-probable frames yielded that, 5 solute-solvent hydrogen bonds are formed with 5-surrounding water molecules for keto and for enol. A similar observation was made for 6, 7 and 8 solute-solvent hydrogen bonds, i.e., X solute-solvent hydrogen bonds are formed with X surrounding water molecules. For further analysis, these frames were extracted with solute and hydrogen bonded water molecules.

4.3.3 Cluster with 5-8 Solvating Water Molecules

The artificial breaking of solvent-solvent hydrogen bonds occurs while extracting clusters containing solute and hydrogen bonded solvent molecules from MD trajectory. Addi-

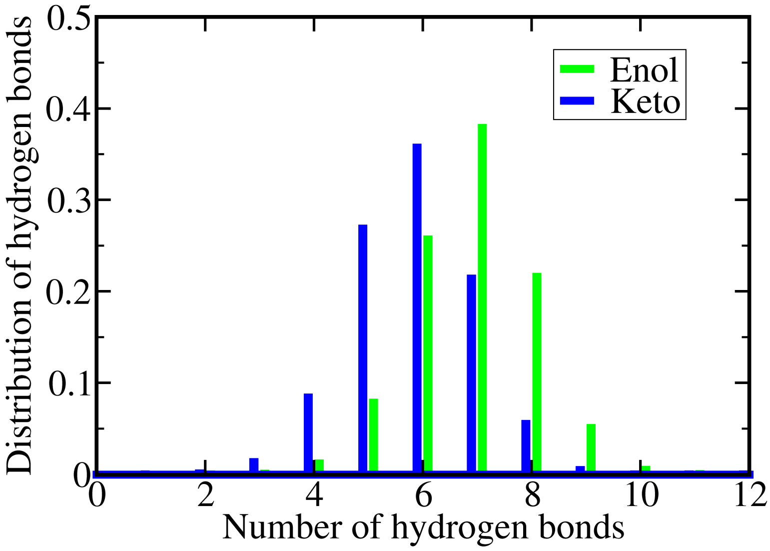


Figure 4.4: Distribution of hydrogen bonds between BA and water molecules. In this distribution plot the total probability of hydrogen bonds is normalized to unity. Total 333334 MD-frames along the production run (100 ns) were used for this analysis.

tionally, orientations of molecules in these clusters are due to kinetic energy obtained in MD simulation. Thus, we performed quantum chemical optimization of geometries obtained for these clusters. As these optimizations are computationally demanding, and thus to perform these calculations for less number of clusters is intended. To do so we further classified these clusters using a number of possible ways a particular number of solute-solvent hydrogen bonds can be formed.

Both keto and enol tautomers have five possible centres for forming hydrogen bonds with water molecules, each of this centre can form a different number of hydrogen bonds. These centres are N_1–H, C_2=O, N_3–H, C_4=O, and C_6=O for keto (C_6–OH for enol). Further, a given number of solute-solvent hydrogen bonds can be formed in many different ways, a different number of hydrogen bonds formed via each centre on BA molecules. This hydrogen distribution is analyzed from a six dimensional histogram. Abscissa for a six dimensional histogram is denoted by generating a string starting with a number (H) followed by six one digit numbers each representing number of hydrogen bonds formed via (N_1–H, C_2=O, N_3–H, C_4=O, C_5–H/H2 and C_6=O/–OH), i.e. $HN_1C_2N_3C_4C_5C_6$. H is the total number of hydrogen bonds formed between solute and solvent molecules. Other numbers denote number of hydrogen bonds formed via each centre. For example a string 7121201 for enol means there are 7 solute-solvent hydrogen bonds and both the N-H groups and the O-H group form 1-hydrogen bond each whereas both the carbonyl groups form 2-hydrogen bonds. Same string for keto tautomer means 7 solute-solvent hydrogen bonds and both N-H and C_6=O groups form one hydrogen bond, while C_2=O and C_4=O forms two hydrogen bonds each. Note that hydrogen bonds formed by C_5–H/H2 are always zero.

Distribution of 5 solute-solvent hydrogen bonds is relatively simple, as for both keto and enol, this distribution peaks, when each centre on BA molecule forms exactly one hydrogen bond with exactly one water molecule (Figure 4.5). It is similar in the case of 8 solute-solvent hydrogen bonds, where 2 hydrogen bonds are formed via each carbonyl/hydroxyl groups and 1 hydrogen bond is formed via each N-H centre (Figure 4.8).

The case is a little complicated for 6 and 7 solute-solvent hydrogen bonds. Distribution of 6 solute-solvent hydrogen bonds within 5-centres on the solute molecules is shown in Figure 4.6. Pie-chart on each centre presents various occurrences of hydrogen bonds with water molecules when total solute-solvent hydrogen bonds were 6, during the course of MD trajectory. Figure 4.7, shows a similar distribution of 7 solute-solvent hydrogen bonds formed in MD trajectory. From this classification, various types of clusters were obtained. Table 4.2 describes types of clusters having some highest percentage based on occurrence in MD trajectory. Clusters with the highest population in each type were chosen for quantum optimization.

Bulk water environment is assumed to be similar for solvated keto and enol tautomers beyond water molecules that are directly hydrogen bonded to solute. In real conditions, water molecules in the clusters form hydrogen bonds with outer water molecules in bulk

Table 4.2: Percent occurrence of clusters chosen for quantum optimizations in the MD trajectory.

Keto		Enol	
Cluster type	Occurrence %[a]	Cluster type	Occurrence %[a]
5111101	69.73	5111101	36.41
6111201/6111102	60.53	6111201	44.04
7121201/7121102	55.14	7121201	56.50
8112202	68.13	8112202	50.84

[a] Clusters have occurrence percentage higher than 50% (except for the case of cluster type for enol 511111 and 611211), which means that any other type of cluster in these cases is not as probable as the ones presented here. For enol 5-hydrogen bonds with water, the second largest abundant cluster is 510121 with percentage 20.73 and for enol 6-hydrogen bonds with water, second largest cluster is 611121 with percentage 21.59.

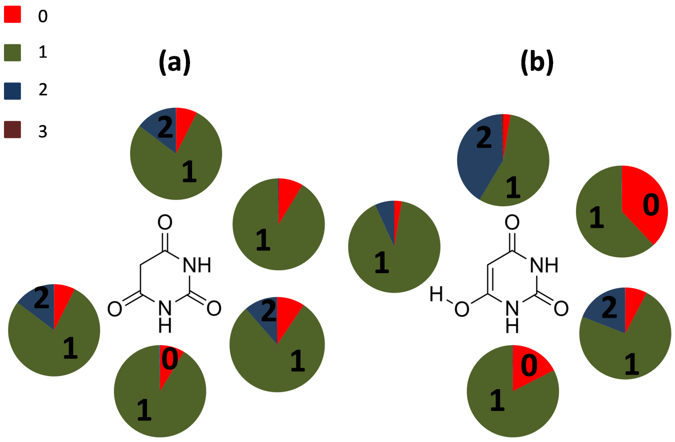


Figure 4.5: Distribution (Pie chart) of 5 solute-solvent hydrogen bonds formed via each hydrogen bond centre on the solute (a) keto and (b) enol. Colours represent number of hydrogen bonds.

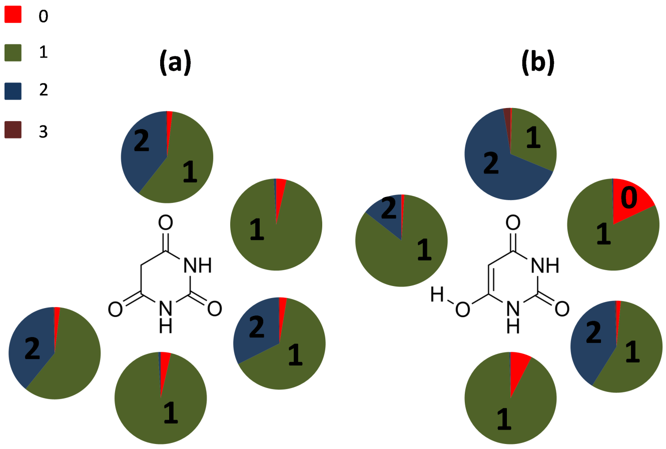


Figure 4.6: Distribution (Pie chart) of 6 solute-solvent hydrogen bonds formed via each hydrogen bond centre on the solute (a) keto and (b) enol. Colours represent number of hydrogen bonds.

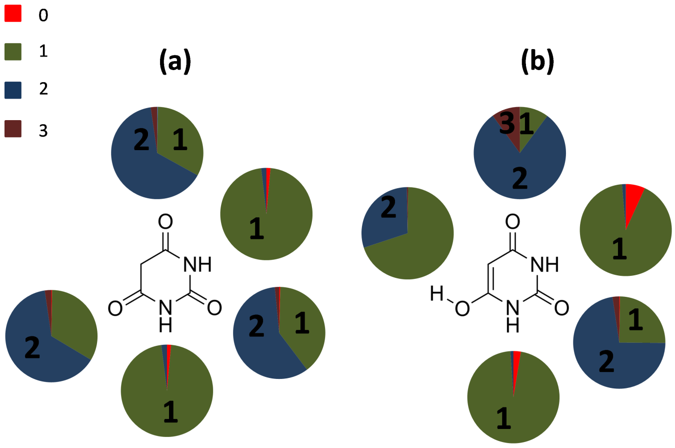


Figure 4.7: Distribution (Pie chart) of 7 solute-solvent hydrogen bonds formed via each hydrogen bond centre on the solute (a) keto and (b) enol. Colours represent number of hydrogen bonds.

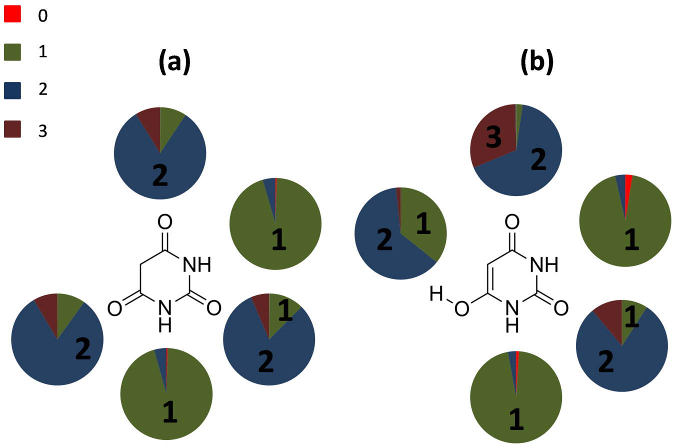


Figure 4.8: Distribution (Pie chart) of 8 solute-solvent hydrogen bonds formed via each hydrogen bond centre on the solute (a) keto and (b) enol. Colours represent number of hydrogen bonds.

water, which are not present in these optimizations, it is an artifact of quantum optimization of clusters. Thus, clusters having alterations in a number of solute-solvent hydrogen bonds after quantum optimizations were omitted from further analysis. Figure 4.9 shows the relative energies of enol-water clusters with respect to keto-water clusters. Data points in red represent the relative energies for clusters obtained by manual optimizations starting with various initial geometries. Relative energies of clusters that were obtained from MD simulations are shown with green data points. We observe that with the addition of water molecules in the clusters not only the relative energy of enol decreases, for clusters containing 8-water molecules the relative energy becomes negative. This observation suggests that enol gets stabilized due to interaction with water molecules and in solution it becomes co-existent with keto tautomer.

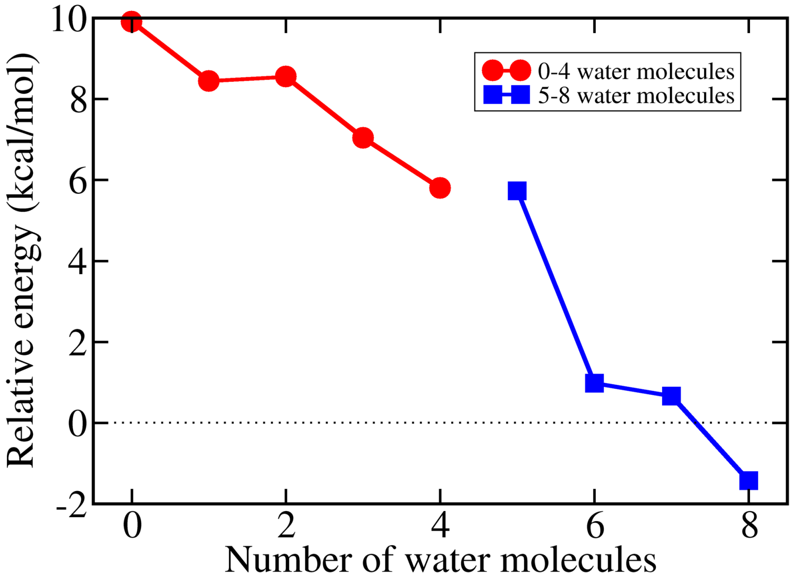


Figure 4.9: The relative stability of enol-water clusters with respect to keto-water clusters. Red indicates the values obtained for clusters with 0-4 water molecules, while green indicates 5-8 water molecules. Clusters with 5-8 water molecules are chosen from the distribution of hydrogen bonds.

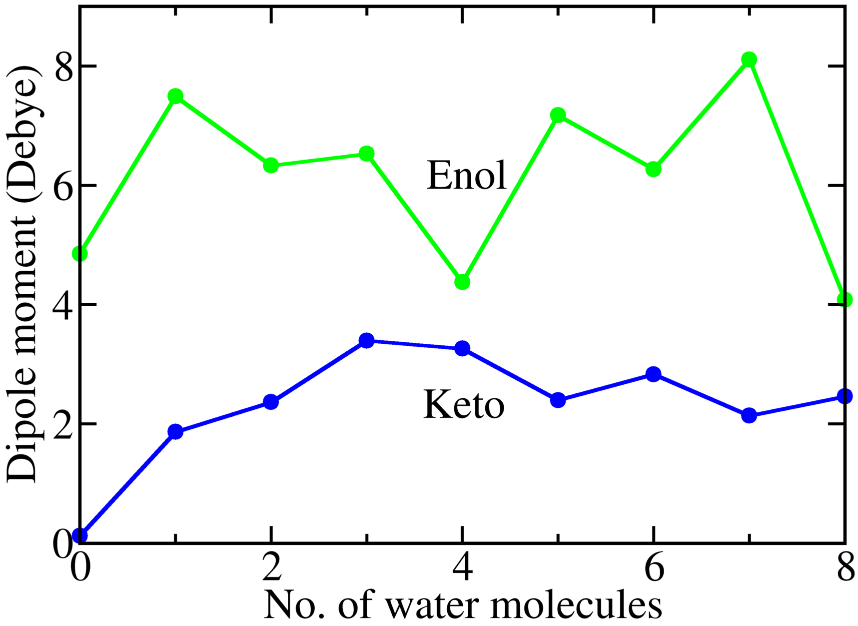


Figure 4.10: Dipole moments of clusters of BA with 0-8 solvating water molecules. Figure shows average dipole moment for clusters with specific number of water molecules.

4.4 Discussion on stabilization of enol tautomer in solution

Figure 4.9 shows that addition of water molecules to clusters containing BA-tautomer results in the increased stabilization of the enol tautomer clusters with respect to the keto tautomer clusters. There are several reasons for this. The first is the higher dipole moment of enol than keto. Figure 4.10 presents the average dipole moments of clusters with keto and enol tautomers. It can be observed that clusters with enol tautomers consistently have higher dipole moment than corresponding clusters with keto tautomer. This suggests that enol would get stabilized more than keto on solvation. Moreover, the MD simulations suggest that on average an extra hydrogen bond is formed by enol tatuomer with water molecules with respect to keto tautomer (see Figure 4.4). Figures 4.5, 4.6, 4.7 and 4.8 show that the enol tautomer has a propensity for the formation of an extra hydrogen bond via C_4=O position than the keto tautomer for the similar position. Additionally charges on hydrogen atoms in enol are higher than in keto as shown in Figure 4.2. It is also observed that oxygen atoms in carbonyl groups at 2 and 4 positions have higher charges in enol than in keto. This hints at stronger hydrogen bond formation, with surrounding water molecules, by enol tautomers, than hydrogen bonds formed by keto tautomer. In summary, stabilization of enol with respect to keto in solution can be explained based on higher dipole moment, tendency for formation of an extra hydrogen bond, and formation of stronger hydrogen bonds due to greater charge separation.

4.5 Conclusions

We have used various theoretical and experimental techniques to study the relative stability of keto and enol tautomers of BA in aqueous phase. Our multipronged approach conclusively suggests a delicate balance between the amount of keto and enol tautomers present as the extent solvation is increased. Theoretical calculations predict that with the increase in the number of water molecules interacting with BA tautomer, the enol tautomer gets more stabilized than the keto tautomer, which is validated by IR spectroscopy (as presented in Appendix B) performed in acetonitrile (which mimics gas-phase conditions), in water and in a mixture of these solvents. Such stabilization of enol tautomer is in-contrast to earlier studies performed with continuum solvation models, which indicates that effect of interactions of explicit water molecules is necessary for correct prediction of relative stability of tautomers.

Theoretical calculations predict that as the extent of solvation is increased to 8 interacting water molecules enol and keto tautomers are likely to coexist due to their small relative energy difference. This suggests that possibly for higher number of water molecules (or in bulk water) also keto tautomers are present along with enol tautomers, though enol becomes more stable than keto in such conditions. This is observed in NMR experiments (as presented in Appendix C), which show very small contribution to the spectra due to keto tautomer in aqueous phase.

In aqueous solution, the enol form gets stabilized due to hydrogen bonding with solvent molecules. There is possibility of one more hydrogen bond in enol form compared to the keto form, and based on charge analysis the bonds are stronger in the enol than the keto tatutomer. This can be contrasted with the stabilization of the enol form in the solid state, which has been previously suggested to be due to inter-molecular hydrogen bonding among BA molecules.

We have proposed a protocol to systematically study the relative stabilization due to solvation i.e., hydrogen bonds, in a collection of tautomers. This protocol can be extended to study other clusters formed by different types of weak interactions like n-π^*, π-π^*, and van Der Waals interactions.

We now point out some possible sources of errors in our calculations. For instance, quantum calculations involve errors due to limitations of the basis sets, density functionals used, and choice of initial structure provided for geometry optimiation. Similarly classical MD simulations can give spurious results as internal (electronic) energy is not described and molecule can deform without these considerations, resulting in unphysical structures. As the theoretical and experimental studies give congruent results in this case, it is possible that the choice theoretical methods are best suit for studying such a problem.

The combined theoretical and experimental study shows that the relatively large energy difference between tautomers of nuelcueobases (11 kcal/mol) can be overturned by solvation. This understand will useful to understand the role of barbituric acid as pre-RNA nucleobase.

Chapter 5

Photophysics of Barbituric Acid

5.1 Introduction

Canonical nucleobases act as chromophores in RNA, with a maximum absorption near 260 nm.[112] Owing their ultrafast nonradiative decay dynamics, these molecules possess a high degree of photostability.[4] This ultrafast internal conversion provides a mechanism by which electronic energy is transferred to molecular vibrations, and then to the surroundings in the form of heat. Various gas-phase experimental and theoretical studies have revealed that the mechanisms behind photostability of nucleobases are sensitive to its tautomeric [113, 114] or isomeric form [115, 116], and substitutions present in the molecule [117, 118].

RNA is a product of chemical and biological evolution. As stated in Chapter 4, proto-RNA, from which current RNA has evolved, have recognition units (nucleobases) different from canonical or contemporary nucleobases.[119] On the surface of early earth, UV radiations were predominant in the ambient environment due to the absence of the ozone layer in the atmosphere. This poses a very important question: What is the role played by photostability of proto-RNA nucleobases, that led to the eventual choices of contemporary nucleobases?

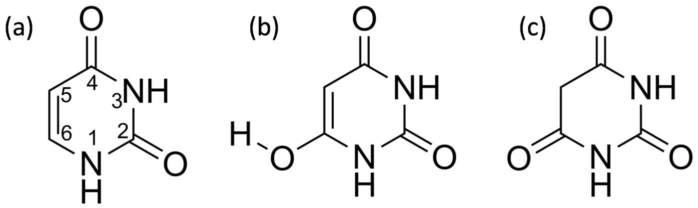


Figure 5.1: Structures of (a) uracil, and BA in (b) enol (c) keto tautomeric forms.

In order to address this question, several researchers have explored the effects of substitution on excited-state processes in current nucleobases.[120, 121] Along with these types of studies, it is also necessary to study the excited-state dynamics of proposed proto-RNA nucleobases and then compare them with those of canonical nucleobases.[111, 122]

As discussed in Chapter 4, the proto-RNA nucleobase barbituric acid (BA) exists predominantly in its enol tautomeric form in solution, while isolated BA is stable in its keto form. Enol form of BA is structurally similar to uracil (Figure 5.1). In a recent study, transient absorption spectroscopy was used to study the spectroscopic and photophysical properties of BA.[111] The steady state absorption spectrum of this molecule was measured for various pH conditions of aqueous buffer. It was suggested that under neutral pH conditions the enol tautomer (5.1b) is predominantly present, while for a lower pH of 2.0, the keto tautomer (5.1c) becomes predominant. Further, the absorption maximum near 257 nm was attributed to the enol tautomer in neutral pH and the keto tautomer in acidic conditions. Excited-state dynamics of this molecule was studied by exciting the molecule at 268 nm. Two timescales of photodecay were observed: 180 and 300 fs. It was found that these lifetimes are very similar in neutral and acidic pH conditions, which led authors to conclude that both enol and keto have very similar ultrafast decay dynamics.

Quantum chemical studies can provide mechanistic insights to ultrafast decay processes. Delchev and Ivanova have studied mechanisms of nonradiative decay of BA, which were published before the experimental work which has been described above.[123] In this work, pathways for photoinduced dissociation of –NH and –OH bonds were calculated for several tautomers of BA. It was found that, $\pi\sigma^*$ and $n\sigma^*$ states play a dominant role in such dissociative mechanisms of nonradiative decay. Later, Shterev and Delchev studied the effect of solvent polarity on the dissociative nonradiative decay mechanism for the keto tautomer of BA (Figure 5.1c) along the –NH stretching mode.[124] It was concluded that in polar solvents the –NH dissociative pathway is disfavoured. Both of these studies were performed at the TDDFT level of theory. However, the experimentally observed ultrafast lifetimes of 180 and 300 fs cannot be explained with these studies.

In this work, we have investigated the deactivation mechanisms of both the tautomers of BA post photoexcitation to lowest bright state. Using the CASSCF method, relevant conical intersections (CI) were optimized and pathways leading to these CIs were calculated using CASPT2 methods. From these pathways we can explain the experimental observation of ultrafast lifetime of the excited states.

5.2 Computational Methods

Ground-state equilibrium geometries of both the tautomers were optimized with the Møller–Plesset second-order perturbation method [22] with Dunning's double-zeta (cc-pVDZ) basis set[58]. The absorption spectra was calculated with vertical excitation energies and oscillator strengths computed using the TDDFT method[125] with the B3LYP functional[99, 100] and 6-31G** basis set. For constructing the absorption spectra, Gaussian band shape, of standard deviation 0.4 eV and normalization constant proportional to the oscillator strength, were assumed. The calculated vertical excitation energies of both the tautomers were shifted by 30 nm for the final calculation of the absorption spectra. The calculations of vertical excitation energies and oscillator strengths were performed with the Gaussian software package.[101]

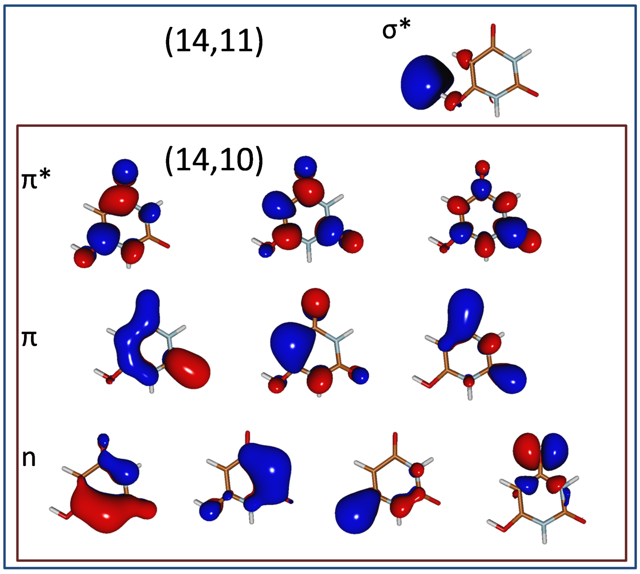


Figure 5.2: Orbitals included in the active space for studying nonradiative decay mechanism of enol tautomer. (14,11) active space consists of (14,10) active space along with σ^* orbital.

We also calculated vertical excitation energies using EOM-CCSD [59, 60] as well as MS-CASPT2 [31] methods. The state averaged (SA) CASSCF calculation, with the lowest 4 states for enol and the lowest 5 states for keto (including ground state), was performed for as reference to the MS-CASPT2 calculation. The active space for enol-tautomer consisted of 3 $\pi - \pi^*$, 4 lone-pair orbitals, and a Rydberg-type σ^* orbital along the hydroxyl group, *i.e.*, a (14,11) active space (Figure 5.2). Calculations for keto-tautomer were performed with a (14,10) active space without including the Rydberg-type orbital (Figure 5.3). These calculations were performed using the 6-31++G** basis set.

For the enol-tautomer, the optimizations of CIs and calculations of pathways, which involve $\pi\sigma^*$ state, were performed with the same active space, state averaging and basis set as used for the calculation of vertical excitation energy. The calculations of excited minimum energy structures, CIs and pathways for nonradiative decay which do not involve the $\pi\sigma^*$ state, were performed with the (14,10) active space, 3 state averaged

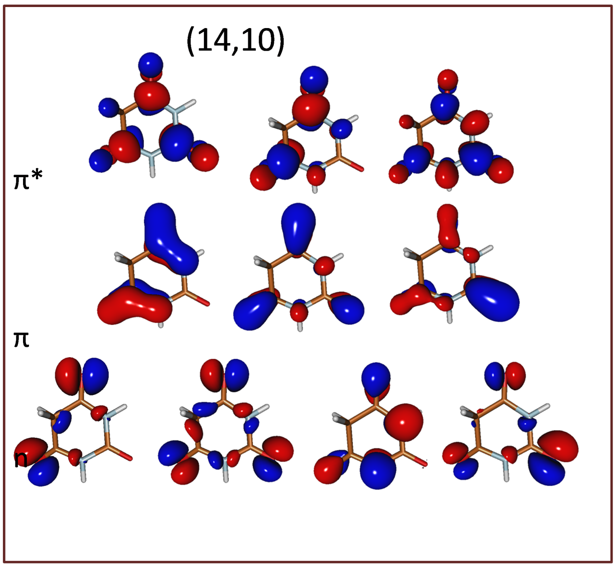


Figure 5.3: Description of (14,10) active space used for studying nonradiative decay mechanism of keto tautomer.

CASSCF wave function, and the 6-31G** basis set.

For keto-tautomer, the calculations of excited state minimum energy structures, CIs and pathways leading to nonradiative decay were calculated using the same active space and state averaging as that used for vertical excitation energy calculations. The 6-31G** basis set was used in these calculations. The state specific (SS) CASPT2 method [30] was used to include dynamical correlation with a level shift of 0.3 au, while calculating pathways for nonradiative decay.

5.3 Equilibrium Geometry and Vertical Excitation Energies

Ground state optimized geometries of keto and enol tautomers are shown in Figure 5.4. The keto tautomer has a slightly non-planar structure, while the enol tautomer has a planar structure. Planarity of the enol tautomer is mainly due to the presence of the $C_5{=}C_6$ double bond, which makes the system aromatic. On absorption of light, the

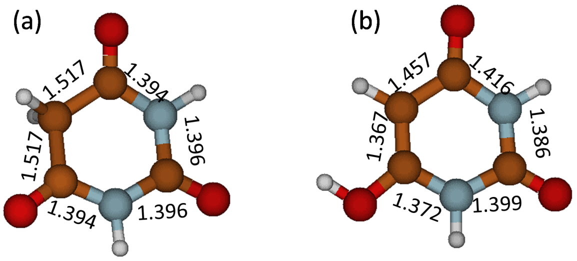


Figure 5.4: Ground state equilibrium geometries of keto and enol tautomers.

Table 5.1: Calculated Vertical Excitation Energies (ΔE) and Oscillator Strengths (f) for keto-tautomer of BA using EOM-CCSD, CASPT2/CASSCF(14,10) active space, and TDDFT methods

	EOM-CCSD/6-31++G**		MS-CASPT2/6-31++G**		CAM-B3LYP/6-31++G**	
Transition	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f
S_1 ($n\pi^*$)	5.17	0.0005	5.48	0.0003	4.99	0.0004
S_2 ($n\pi^*$)	5.48	0.0020	5.70	0.0011	5.35	0.0016
S_3 ($n\pi^*$)	6.31	0.0039	6.49	0.0004	6.08	0.0024
S_4 ($\pi\pi^*$)	6.72	0.2210	6.97	0.3250	6.48	0.1499

electronic excitation is supposed to take place from these geometries.

Table 5.1 and Table 5.2 list the calculated vertical excitation energies and oscillator strengths of states up to the lowest bright state for keto and enol tautomers, respectively. Both keto and enol have a $\pi\pi^*$ state as their bright state. A bright state is identified by the corresponding oscillator strength. The inspection of excitation energies suggests that keto absorbs at higher energy (smaller wavelength) than enol. The difference in the vertical excitation energies of keto and enol can be attributed to the absence/presence of $C_5{=}C_6$ double bond. As this bond is present in enol tautomer, its excitation energies can be compared with the similar-in-nature pyrimidine nucleobase, uracil.

For the enol tautomer, $\pi\pi^*$ excitation energies, 5.75 eV (EOM-CCSD), 6.20 eV (MS-CASPT2) and 5.57 eV (TDDFT), are comparable to that of uracil as presented in Ref [126], with values of 6.24 eV and 5.79 eV calculated using two different active spaces

Table 5.2: Calculated Vertical Excitation Energies (ΔE) and Oscillator Strengths (f) for enol-tautomer of BA using EOM-CCSD, CASPT2/CASSCF(14,11) active space, and TDDFT methods

	EOM-CCSD/6-31++G**		MS-CASPT2/6-31++G**		CAM-B3LYP/6-31++G**	
Transition	ΔE (eV)	f	ΔE (eV)	f	ΔE (eV)	f
S_1 $n\pi^*$	5.48	0.0001	5.46	0.0002	5.33	0.0000
S_2 $\pi\sigma^*$	5.74	0.0043	6.15	0.0093	5.66	0.0093
S_3 $\pi\pi^*$	5.75	0.2751	6.20	0.5324	5.57	0.1965

with the MRCI method. This similarity in excitation energies is expected due to the structural similarity between the enol tautomer and uracil. Further, absorption spectra of uracil in various solvents is observed to be near 260 nm.[127] This comparison hints to the point that the maximum in the steady state absorption absorption [111] near 260 nm is due to the enol-tautomer.

We calculated the absorption spectrum, using TDDFT(B3LYP)/6-31++G** method, for BA, assuming a variable relative concentration of keto and enol tautomers, as presented in Figure 5.5. The abscissa in this figure is shifted by 30 nm for a better match with experimental values. Such a systematic shift is reasonable as zero-point energy correction and solvent effects were not considered while calculating the spectra. This calculation suggests that only the enol tautomer absorbs the radiations near 260 nm.

Transient absorption experiments for BA, as presented in Ref [111], were performed by exciting the molecule at 267 nm. Based on the assumption that keto and enol, both absorb near 260 nm, and keto being more stable in acidic pH while enol in neutral conditions, the authors concluded that keto and enol have very similar excited-state dynamics. However, from the calculated vertical excitation energies and simulated absorption spectra, we believe that under the described experimental conditions, only the enol tautomer was excited in both pH conditions. We hypothesize that this was the reason why similar photodecay dynamics was observed, and not the authors' conclusion that both keto and enol have similar decay dynamics. We further strengthen this hypothesis by show-

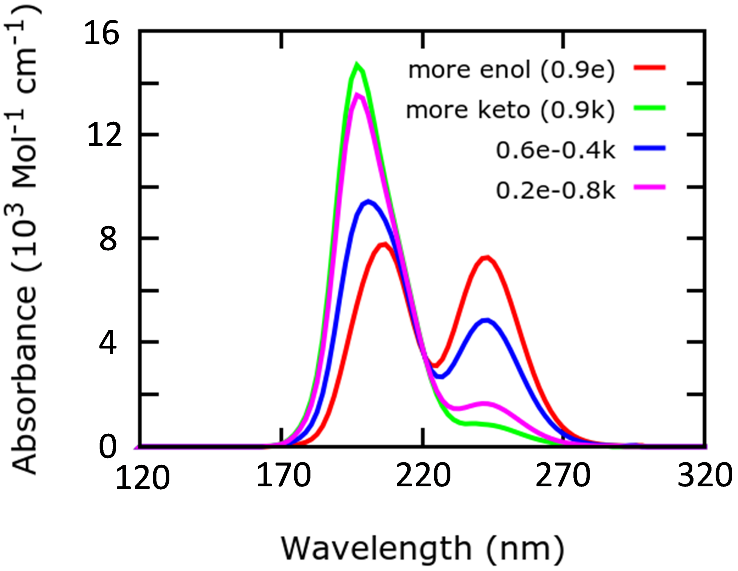


Figure 5.5: Calculated absorption spectra of BA. By assuming a relative concentration of keto and enol tautomers, absorption spectra is calculated. Total absorbance is the sum of the product of absorbance of a tautomer and it's relative concentration.

ing, using quantum chemical calculations, that keto and enol have different photodecay pathways.

5.4 Nonradiative Decay of Keto-tautomer

Following photoexcitation, the molecule reaches the first bright electronic state, on which we study the nonradiative decay mechanism. First we tried to optimize the minimum energy geometry for this lowest bright ($\pi\pi^*$) state, as it is natural to assume that the molecule will relax to this geometry. However, the equilibrium geometry for this state could not be located, as the state starts mixing with the lower states, as a result of which our calculations did not converge. Therefore, we looked for CIs between the $\pi\pi^*$ state and lower $n\pi^*$ states. These calculations did not converge either. For the same reason, a CI between the $\pi\pi^*$ and the ground state could not be optimized, even though

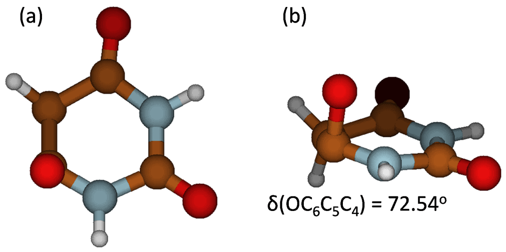


Figure 5.6: 5 SA-CASSCF optimized geometry of S_1/S_0 CI (a) top and (b) side view.

calculations were started with various geometries.

It is known that twisting of a double bond stabilizes $\pi\pi^*$ and $n\pi^*$ states, as the overlap between the corresponding π^* orbital increases. Therefore, we looked for CIs with the ground state by twisting along the C_4–O bending coordinate. In this manner, we were able to optimize the S_1/S_0 CI, as shown in Figure 5.6. This CI shows out-of-plane displacement of oxygen. A possible pathway for nonradiative decay from the S_1 state was constructed by connecting the Franck-Condon (FC) geometry to the CI using geometries generated by linearly interpolated internal coordinates (LIIC), as shown in Figure 5.7. As there is no direct CI from the initially excited $\pi\pi^*$ state to lower states, it is possible that the nonradiative decay in the keto tautomer is a comparatively slow process.

5.5 Nonradiative Decay of Enol-tautomer

Nonradiative decay of the enol tautomer was studied by first optimizing the initially populated $\pi\pi^*$ state equilibrium geometry. The optimization calculation quickly reached a crossing with a lower $\pi\sigma^*$ state, and we were able to optimize a CI between these two states (Figure 5.8a). This structure is very similar to the FC geometry. Figure 5.9a represents the energy profiles of the excited and ground electronic states along the LIIC pathway connecting the FC structure to the $\pi\pi^*/\pi\sigma^*$ CI. From this geometry, the

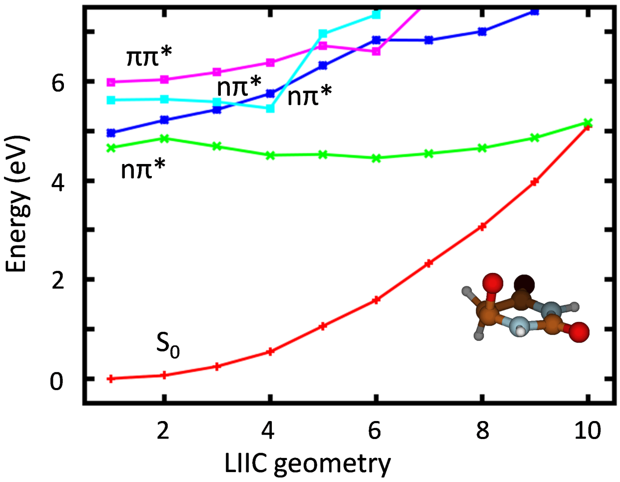


Figure 5.7: Profile of excited and ground state energies calculated using SS-CASPT2 method along LIIC pathway for connecting FC geometry to S_1/S_0 CI.

molecular wavepacket transfers to the $\pi\sigma^*$ state. Further, we optimized the $\pi\sigma^*/S_0$ CI along the O–H stretching coordinate as shown in Figure 5.8b. The $\pi\sigma^*/S_0$ CI, is planar structure and has an O–H distance of 1.75 Å. Relaxed scan was performed for the $\pi\sigma^*$ state along the O–H stretching coordinate by relaxing all other coordinates for a fixed value of O–H distance. Figure 5.9b shows the energy profiles calculated using the SS-CASPT2 method. This pathway was found to have no barrier for accessing the $\pi\pi^*/\pi\sigma^*$ CI along the $\pi\sigma^*$ state.

We also optimized a CI between the $\pi\pi^*$ and the ground state along the C_5=C_6 twisting coordinate, as was observed for uracil.[126] This CI involves out-of-plane distortion of the hydroxyl group as well as the hydrogen at C_5 position (Figure 5.10). Figure 5.11 shows the SS-CASPT2 energy profiles along the LIIC pathway connecting the FC geometry to the $\pi\pi^*/S_0$ CI. Similar to uracil, this pathway is barrierless in the case of enol.

We are now in a position to assign the lifetimes obtained in the experiment.[111] The fast lifetime of 180 fs is assigned to the direct transfer of the population from the

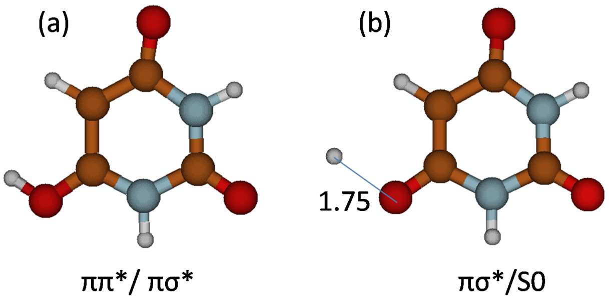


Figure 5.8: 4 SA-CASSCF optimized structures for (a) $\pi\pi^*/\pi\sigma^*$ CI and (b) $\pi\sigma^*/S_0$ CI.

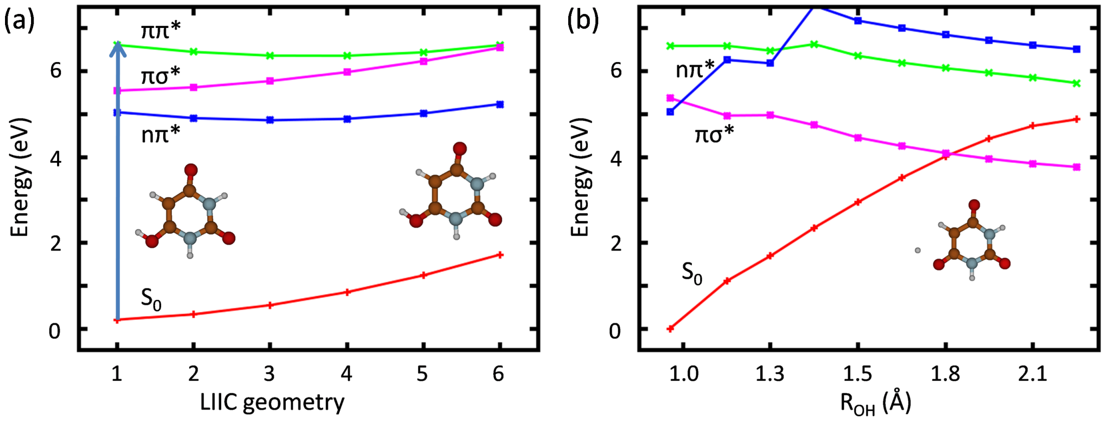


Figure 5.9: SS-CASPT2 profiles of electronic energies for (a) LIIC pathway connecting FC geometry to $\pi\pi^*/\pi\sigma^*$ CI and (b) relaxed scan for $\pi\sigma^*$ state O–H stretching coordinate.

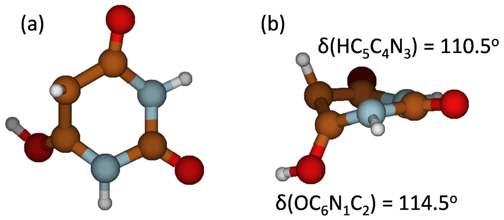


Figure 5.10: $\pi\pi^*/S_0$ CI optimized using 3 SA-CASSCF method: (a) top (b) side view.

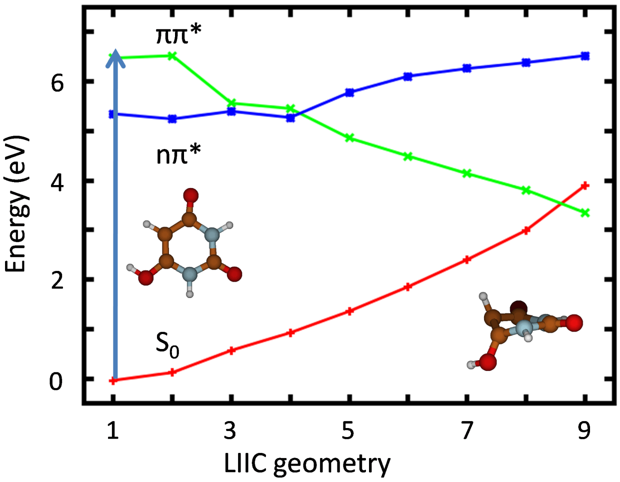


Figure 5.11: LIIC pathway for connecting FC geometry to $\pi\pi^*/\mathrm{S}_0$ CI.

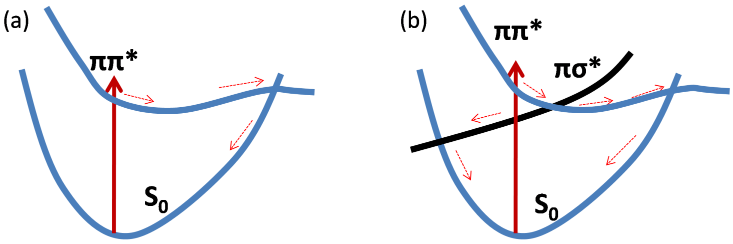


Figure 5.12: Schematic of possible mechanisms of nonradiative decay in (a) uracil and (b) enol tautomer of BA.

$\pi\pi^*$ to the ground state via the $\pi\pi^*/\mathrm{S}_0$ CI. The comparatively slower lifetime of 300 fs is assigned to the nonradiative decay to the ground state via the $\pi\pi^*/\pi\sigma^*$ CI along the –OH stretching coordinate. It is evident from these calculations that the keto and enol tautomers have very different photodecay dynamics.

5.6 Comparison with Nonradiative Decay of Uracil

Excited state dynamics of uracil has been studied using various pump-probe techniques. Transient-absorption experiments found that uracil has a picosecond (ps) lifetime for nonradiative decay.[4][128] Using 267 nm excitation and ionization probe two lifetimes of 0.13 ps and 1.05 ps were observed.[129] Another study performed with a similar method found lifetimes of 0.07–0.09 ps and 2–3 ps.[130] These results suggest that uracil has at least one lifetime > 1 ps. Recently, it was proposed that this higher lifetime in uracil is due to excited molecule wavepacket which gets trapped in the $\pi\pi^*$ minimum region.[131] The schematic proposed in this study is shown in Figure 5.12. In the case of BA however, transient absorption spectroscopic experiment gave two lifetimes for nonradiative decay: 180 fs and 300 fs.[111]. In our work, we have attributed this to a $\pi\pi^*/\mathrm{S}_0$ CI and a $\pi\pi^*/\pi\sigma^*$ CI, respectively.

This implies that the main difference between the enol tautomer and uracil is the presence of a nonradiative decay pathway involving a $\pi\sigma^*$ state in the case of enol. While this may not seem that stark a difference, we need to consider what a pathway involving a $\pi\sigma^*$ state means. The $\pi\sigma^*$ state is a dissociative state, meaning that this excited state is stabilized upon elongation of the σ bond. While this can lead to the molecule returning to the ground state via a crossing with it, it is also possible for the wavepacket to continue along the $\pi\sigma^*$ state causing dissociation of the proton from the molecule. This possible dissociation pathway takes away from a molecule's ability to be photostable. In fact, the presence of a pathway along the $\pi\sigma^*$ state in the enol tautomer of BA could possibly be the reason why it is not a contemporary nucleobase.

Chapter 6

Discussion

Our studies on the mechanisms for photostability of DHI and BA are presented in this thesis. These mechanisms were studied using multi-reference methods for isolated molecules. Although biomolecules are present in a complex environment, in-depth mechanistic study photostability of isolated molecules provides detailed picture of innate photophysical properties of the molecules. Such theoretical studies often useful in providing mechanistic understandings of experimental observations. Particular to this thesis, our studies on photophysics of DHI and BA are able to explain the experimental observations of excited state lifetimes.

Multi-reference methods are necessary for studying mechanistic photostability for two reasons: one these mechanisms involve crossing of electronic states and two constructing nonradiative decay pathways include breaking/making of bonds. As in the case of DHI, presence of barrier along the nonradiative decay pathway was not captured in earlier study performed using single-reference methods. Further, it is found that the presence of barrier is related to change in character of the electronic state.

$\pi\sigma^*$ states are known to play a significant role in governing nonradiative decay of heteroaromatic molecules. Photophysics of DHI as well as of enol tautomer of BA is influenced by the presence of $\pi\sigma^*$ states. For a correct description of $\pi\sigma^*$ states, diffuse functions must be included in the basis set.

Hydrogen bonding destabilizes the $\pi\sigma^*$ state, if Rydberg-type σ^* orbital is situated on the donor atom. This observation is explained due to weak interaction between non-bonding orbital on acceptor atom and σ^* orbital on donor atom.

Nonradiative decay along $\pi\sigma^*$ state is attained via stretching of X–H (X:heteroatom) bond. Along the stretched X–H bond, $\pi\sigma^*$ state has a conical intersection (CI) with the ground state. As this CI, molecular wavepacket can branch into three directions: one relaxing to the ground state minimum, the other two possibilities are dissociation and formation of excited or ground state products. Dynamical calculations around such CIs are necessary for building a complete picture of possible outcomes after CI is accessed.

We have also studied relative stability of keto and enol tautomers of BA, in this work solvent effects were treated explicitly. A protocol is proposed for choosing statistically meaningful structure, as presented in Appendix 6. This protocol will be useful for future studies dealing with such problems.

List of Publications

Publications included in this thesis

1. **Datar, A.;** Hazra, A. "Pathways for Excited-State Nonradiative Decay of 5,6-Dihydroxyindole, a Building Block of Eumelanin." *J. Phys. Chem. A,* **2017,** *121,* 2790-2797.

2. **Datar, A.;** Paithankar, H.; Deb, P.; Chugh, J.; Bagchi, S.; Mukherjee, A.; Hazra, A. "Solvation-Driven Keto-to-Enol Tautomerization of Barbituric Acid." *Manuscript under preparation*

3. **Datar, A.;** Hazra, A. "Pathways for Excited-State Nonradiative Decay of Barbituric Acid, a Proto-RNA Nucleobase." *Manuscript under preparation*

Publications not included in this thesis

1. Kand, D.; Mandal, P. S.; **Datar, A.;** Talukdar, P. "Iminocoumarin based fluorophores: indispensable scaffolds for rapid, selective and sensitive detection of thiophenol" *Dyes and Pigments* **2014,** *106,* 25-31.

2. Roy, A.; **Datar, A.;** Kand, D.; Talukdar, P. "A fluorescent off–on NBD-probe for F− sensing: theoretical validation and experimental studies." *Organic and biomolecular chemistry* **2014,** *12,* 2143-2149.

Bibliography

[1] M. Klessinger and Michl J. *Excited States and Photochemistry of Organic Molecules.* New York: VCH Publishers, Inc., 1995.

[2] Jakob Klán, Petr and Wirz. *Photochemistry of organic compounds : from concepts to practice.* Hoboken: John Wiley & Sons, 2009.

[3] Nikiforos Kollias et al. “New trends in photobiology: Photoprotection by melanin”. In: *Journal of Photochemistry and Photobiology B: Biology* 9.2 (1991), pp. 135 –160. DOI: `10.1016/1011-1344(91)80147-A`.

[4] Carlos E. Crespo-Hernndez et al. “Ultrafast Excited-State Dynamics in Nucleic Acids”. In: *Chemical Reviews* 104.4 (2004), pp. 1977–2020. DOI: `10.1021/cr0206770`.

[5] Chao-Hsun YANG et al. “Comparative Study on the Photostability of Arbutin and Deoxy Arbutin: Sensitivity to Ultraviolet Radiation and Enhanced Photostability by the Water-Soluble Sunscreen, Benzophenone-4”. In: *Bioscience, Biotechnology, and Biochemistry* 77.5 (2013), pp. 1127–1130. DOI: `10.1271/bbb.130042`.

[6] Jacob C. Dean et al. “Plant Sunscreens in the UV-B: Ultraviolet Spectroscopy of Jet-Cooled Sinapoyl Malate, Sinapic Acid, and Sinapate Ester Derivatives”. In: *Journal of the American Chemical Society* 136.42 (2014), pp. 14780–14795. DOI: `10.1021/ja5059026`.

[7] Charles S. Cockell and Gerda Horneck. “The History of the UV Radiation Climate of the EarthTheoretical and Space-based Observations”. In: *Photochem-*

istry and Photobiology 73.4 (), pp. 447–451. DOI: `10.1562/0031-8655(2001)0730447THOTUR2.0.CO2`.

[8] Ingrid Cnossen et al. “Habitat of early life: Solar X-ray and UV radiation at Earth’s surface 43.5 billion years ago”. In: *Journal of Geophysical Research: Planets* 112.E2 (). DOI: `10.1029/2006JE002784`.

[9] Carl Sagan. “Ultraviolet selection pressure on the earliest organisms”. In: *Journal of Theoretical Biology* 39.1 (1973), pp. 195 –200. DOI: `10.1016/0022-5193(73)90216-6`.

[10] Howard E. Zimmerman. “Molecular Orbital Correlation Diagrams, Mobius Systems, and Factors Controlling Ground- and Excited-State Reactions. II”. In: *Journal of the American Chemical Society* 88.7 (1966), pp. 1566–1567. DOI: `10.1021/ja00959a053`.

[11] Howard E. Zimmerman. “On Molecular Orbital Correlation Diagrams, the Occurrence of Mbius Systems in Cyclization Reactions, and Factors Controlling Ground- and Excited-State Reactions. I”. In: *Journal of the American Chemical Society* 88.7 (1966), pp. 1564–1565. DOI: `10.1021/ja00959a052`.

[12] J. Michl. “Photochemical reactions of large molecules. I. A simple physical model of photochemical reactivity”. In: *Molecular Photochemistry* 4 (), pp. 243–255.

[13] E. Teller. “Internal Conversion in Polyatomic Molecules”. In: *Israel Journal of Chemistry* 7.2 (), pp. 227–235. DOI: `10.1002/ijch.196900034`.

[14] Fernando Bernardi, Massimo Olivucci, and Michael A. Robb. “Potential energy surface crossings in organic photochemistry”. In: *Chem. Soc. Rev.* 25 (5 1996), pp. 321–328. DOI: `10.1039/CS9962500321`.

[15] T Helgaker, P Jørgensen, and J Olsen. *Molecular Electronic Structure Theory.* Chichester: John Wiley & Sons, LTD, 2000.

[16] M Born and R Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: `10.1002/andp.19273892002`.

[17] J Von Neumann and E Wigner. "No crossing rule". In: *Z. Phys* 30 (1929), pp. 467–470.

[18] Fernando Bernardi, Massimo Olivucci, and Michael A. Robb. "Potential energy surface crossings in organic photochemistry". In: *Chem. Soc. Rev.* 25 (5 1996), pp. 321–328. DOI: `10.1039/CS9962500321`.

[19] David R. Yarkony. "Conical Intersections: The New Conventional Wisdom". In: *The Journal of Physical Chemistry A* 105.26 (2001), pp. 6277–6293. DOI: `10.1021/jp003731u`.

[20] Andrzej L Sobolewski and Wolfgang Domcke. "Conical intersections induced by repulsive $\pi\sigma^*$ states in planar organic molecules: malonaldehyde, pyrrole and chlorobenzene as photochemical model systems". In: *Chemical Physics* 259.2 (2000), pp. 181 –191. DOI: `10.1016/S0301-0104(00)00203-2`.

[21] Mario Barbatti et al. "Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases". In: *Proceedings of the National Academy of Sciences* 107.50 (2010), pp. 21453–21458. DOI: `10.1073/pnas.1014982107`.

[22] Chr. Møller and M. S. Plesset. "Note on an Approximation Treatment for Many-Electron Systems". In: *Phys. Rev.* 46 (7 1934), pp. 618–622. DOI: `10.1103/PhysRev.46.618`.

[23] George D. Purvis and Rodney J. Bartlett. "A full coupledcluster singles and doubles model: The inclusion of disconnected triples". In: *The Journal of Chemical Physics* 76.4 (1982), pp. 1910–1918. DOI: `10.1063/1.443164`.

[24] Gustavo E. Scuseria, Curtis L. Janssen, and Henry F. Schaefer. "An efficient reformulation of the closedshell coupled cluster single and double excitation (CCSD)

equations". In: *The Journal of Chemical Physics* 89.12 (1988), pp. 7382–7387. DOI: `10.1063/1.455269`.

[25] Bjrn O. Roos. "The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations". In: *Advances in Chemical Physics*. Wiley-Blackwell, 2007, pp. 399–445. DOI: `10.1002/9780470142943.ch7`.

[26] Bjrn O. Roos, Peter R. Taylor, and Per E.M. Sigbahn. "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach". In: *Chemical Physics* 48.2 (1980), pp. 157 –173. DOI: `10.1016/0301-0104(80)80045-0`.

[27] Pter G. Szalay et al. "Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications". In: *Chemical Reviews* 112.1 (2012), pp. 108–181. DOI: `10.1021/cr200137a`.

[28] Bjrn O. Roos et al. "A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a CASSCF reference wavefunction". In: *Chemical Physics* 66.1 (1982), pp. 197 –207. DOI: `10.1016/0301-0104(82)88019-1`.

[29] Kerstin. Andersson et al. "Second-order perturbation theory with a CASSCF reference function". In: *The Journal of Physical Chemistry* 94.14 (1990), pp. 5483–5488. DOI: `10.1021/j100377a012`.

[30] Kerstin Andersson, Perke Malmqvist, and Bjrn O. Roos. "Secondorder perturbation theory with a complete active space selfconsistent field reference function". In: *The Journal of Chemical Physics* 96.2 (1992), pp. 1218–1226. DOI: `10.1063/1.462209`.

[31] James Finley et al. "The multi-state CASPT2 method". In: *Chemical Physics Letters* 288.2 (1998), pp. 299 –306. DOI: `10.1016/S0009-2614(98)00252-8`.

[32] Bjrn O. Roos et al. “Applications of level shift corrected perturbation theory in electronic spectroscopy”. In: *Journal of Molecular Structure: THEOCHEM* 388 (1996), pp. 257 –276. DOI: `10.1016/S0166-1280(96)80039-X`.

[33] Luis Serrano-Andrs and Manuela Merchn. “Quantum chemistry of the excited state: 2005 overview”. In: *Journal of Molecular Structure: THEOCHEM* 729.1 (2005). Proceedings of the 30th International Congress of Theoretical Chemists of Latin Expression, pp. 99 –108. DOI: `10.1016/j.theochem.2005.03.020`.

[34] H.-J. Werner et al. *MOLPRO, version 2012.1, a package of ab initio programs.* 2012.

[35] Hans-Joachim Werner et al. “Molpro: a general-purpose quantum chemistry program package”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2 (2012), pp. 242–253. DOI: `10.1002/wcms.82`.

[36] G. Prota. *Melanins and Melanogenesis.* San Diego, CA: Academic Press, 1992.

[37] Paul Meredith et al. “Towards structurepropertyfunction relationships for eumelanin”. In: *Soft Matter* 2 (2006), pp. 37–44. DOI: `10.1039/B511922G`.

[38] Marco d’Ischia et al. “Chemical and Structural Diversity in Eumelanins: Unexplored Bio-Optoelectronic Materials”. In: *Angewandte Chemie International Edition* 48.22 (), pp. 3914–3921. DOI: `10.1002/anie.200803786`.

[39] J. Brian Nofsinger, Tong Ye, and John D. Simon. “Ultrafast Nonradiative Relaxation Dynamics of Eumelanin”. In: *The Journal of Physical Chemistry B* 105.14 (2001), pp. 2864–2866. DOI: `10.1021/jp004045y`.

[40] Sheng Meng and Efthimios Kaxiras. “Theoretical Models of Eumelanin Protomolecules and their Optical Properties”. In: *Biophysical Journal* 94.6 (2008), pp. 2095 –2105. DOI: `10.1529/biophysj.107.121087`.

[41] Magdalena Gauden et al. “Ultrafast Excited State Dynamics of 5,6-Dihydroxyindole, A Key Eumelanin Building Block: Nonradiative Decay Mechanism”. In: *The Journal of Physical Chemistry B* 113.37 (2009), pp. 12575–12580. DOI: `10.1021/jp903190k`.

[42] Annemarie Huijser, Alessandro Pezzella, and Villy Sundstrm. “Functionality of epidermal melanin pigments: current knowledge on UV-dissipative mechanisms and research perspectives”. In: *Phys. Chem. Chem. Phys.* 13 (20 2011), pp. 9119–9127. DOI: `10.1039/C1CP20131J`.

[43] L. E. Bolvar-Marinez, D. S. Galvo, and M. J. Caldas. “Geometric and Spectroscopic Study of Some Molecules Related to Eumelanins. 1. Monomers”. In: *The Journal of Physical Chemistry B* 103.15 (1999), pp. 2993–3000. DOI: `10.1021/jp983336s`.

[44] Klaus B. Stark et al. “Spectroscopic Study and Simulation from Recent Structural Models for Eumelanin: I. Monomer, Dimers”. In: *The Journal of Physical Chemistry B* 107.13 (2003), pp. 3061–3067. DOI: `10.1021/jp0266594`.

[45] Yuri V. Il’ichev and John D. Simon. “Building Blocks of Eumelanin: Relative Stability and Excitation Energies of Tautomers of 5,6-Dihydroxyindole and 5,6-Indolequinone”. In: *The Journal of Physical Chemistry B* 107.29 (2003), pp. 7162–7171. DOI: `10.1021/jp034702x`.

[46] B. J. Powell et al. “A first-principles density-functional calculation of the electronic and vibrational structure of the key melanin monomers”. In: *The Journal of Chemical Physics* 120.18 (2004), pp. 8608–8615. DOI: `10.1063/1.1690758`.

[47] Efthimios Kaxiras et al. “Structural Model of Eumelanin”. In: *Phys. Rev. Lett.* 97 (21 2006), p. 218102. DOI: `10.1103/PhysRevLett.97.218102`.

[48] Andrzej L. Sobolewski and Wolfgang Domcke. “Photophysics of Eumelanin: Ab Initio Studies on the Electronic Spectroscopy and Photochemistry of 5,6-Dihydroxyindole”. In: *ChemPhysChem* 8.5 (), pp. 756–762. DOI: `10.1002/cphc.200600768`.

[49] Giacomo Prampolini, Ivo Cacelli, and Alessandro Ferretti. “Intermolecular interactions in eumelanins: a computational bottom-up approach. I. small building blocks”. In: *RSC Adv.* 5 (48 2015), pp. 38513–38526. DOI: `10.1039/C5RA03773E`.

[50] Deniz Tuna et al. “Onset of the Electronic Absorption Spectra of Isolated and -Stacked Oligomers of 5,6-Dihydroxyindole: An Ab Initio Study of the Building Blocks of Eumelanin”. In: *The Journal of Physical Chemistry B* 120.14 (2016), pp. 3493–3502. DOI: `10.1021/acs.jpcb.6b01793`.

[51] Mukunda Mandal et al. “Feasibility of Ionization-Mediated Pathway for Ultraviolet-Induced Melanin Damage”. In: *The Journal of Physical Chemistry B* 119.42 (2015), pp. 13288–13293. DOI: `10.1021/acs.jpcb.5b08750`.

[52] Per E. M. Siegbahn et al. “The complete active space SCF (CASSCF) method in a NewtonRaphson formulation with application to the HNO molecule”. In: *The Journal of Chemical Physics* 74.4 (1981), pp. 2384–2396. DOI: `10.1063/1.441359`.

[53] Kerstin Andersson, Perke Malmqvist, and Bjrn O. Roos. “Secondorder perturbation theory with a complete active space selfconsistent field reference function”. In: *The Journal of Chemical Physics* 96.2 (1992), pp. 1218–1226. DOI: `10.1063/1.462209`.

[54] Andrzej L Sobolewski and Wolfgang Domcke. “Conical intersections induced by repulsive 1* states in planar organic molecules: malonaldehyde, pyrrole and chlorobenzene as photochemical model systems”. In: *Chemical Physics* 259.2 (2000), pp. 181 –191. DOI: `10.1016/S0301-0104(00)00203-2`.

[55] Michael N. R. Ashfold et al. “* excited states in molecular photochemistry”. In: *Phys. Chem. Chem. Phys.* 12 (6 2010), pp. 1218–1238. DOI: `10.1039/B921706A`.

[56] Gareth M. Roberts and Vasilios G. Stavros. “The role of * states in the photochemistry of heteroaromatic biomolecules and their subunits: insights from gas-phase femtosecond spectroscopy”. In: *Chem. Sci.* 5 (5 2014), pp. 1698–1722. DOI: `10.1039/C3SC53175A`.

[57] Hanna Reisler and Anna I. Krylov. “Interacting Rydberg and valence states in radicals and molecules: experimental and theoretical studies”. In: *International Reviews in Physical Chemistry* 28.2 (2009), pp. 267–308. DOI: `10.1080/01442350902989170`.

[58] Thom H. Dunning. “Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen”. In: *The Journal of Chemical Physics* 90.2 (1989), pp. 1007–1023. DOI: `10.1063/1.456153`.

[59] Henrik Koch and Poul Jo/rgensen. “Coupled cluster response functions”. In: *The Journal of Chemical Physics* 93.5 (1990), pp. 3333–3344. DOI: `10.1063/1.458814`.

[60] John F. Stanton and Rodney J. Bartlett. “The equation of motion coupledcluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties”. In: *The Journal of Chemical Physics* 98.9 (1993), pp. 7029–7039. DOI: `10.1063/1.464746`.

[61] R. Ditchfield, W. J. Hehre, and J. A. Pople. “SelfConsistent MolecularOrbital Methods. IX. An Extended GaussianType Basis for MolecularOrbital Studies of Organic Molecules”. In: *The Journal of Chemical Physics* 54.2 (1971), pp. 724–728. DOI: `10.1063/1.1674902`.

[62] Serhiy Perun, Andrzej L. Sobolewski, and Wolfgang Domcke. “Ab Initio Studies on the Radiationless Decay Mechanisms of the Lowest Excited Singlet States of 9H-Adenine”. In: *Journal of the American Chemical Society* 127.17 (2005), pp. 6257–6265. DOI: `10.1021/ja044321c`.

[63] Angelo Giussani et al. “Essential on the Photophysics and Photochemistry of the Indole Chromophore by Using a Totally Unconstrained Theoretical Approach”. In: *Journal of Chemical Theory and Computation* 7.12 (2011), pp. 4088–4096. DOI: `10.1021/ct200646r`.

[64] Angelo Giussani et al. “The Supporting Information of: Essential on the Photophysics and Photochemistry of the Indole Chromophore by Using a Totally Unconstrained Theoretical Approach”. In: *Journal of Chemical Theory and Computation* 7.12 (2011), pp. 4088–4096. DOI: `10.1021/ct200646r`.

[65] A. L. Sobolewski et al. “Excited-state hydrogen detachment and hydrogen transfer driven by repulsive 1* states: A new paradigm for nonradiative decay in aromatic biomolecules”. In: *Phys. Chem. Chem. Phys.* 4 (7 2002), pp. 1093–1100. DOI: `10.1039/B110941N`.

[66] Ruth Livingstone et al. “Following the excited state relaxation dynamics of indole and 5-hydroxyindole using time-resolved photoelectron spectroscopy”. In: *The Journal of Chemical Physics* 135.19 (2011), p. 194307. DOI: `10.1063/1.3659231`.

[67] B. J. Cafferty et al. “Spontaneous Formation and Base Pairing of Plausible Prebiotic Nucleotides in Water”. In: *Nat. Commun.* 7 (2016), p. 11328. DOI: `DOI: 10.1038/ncomms11328`.

[68] N. V. Hud et al. “The Origin of RNA and ”My Grandfather’s Axe””. In: *Chem. Biol.* 20 (2013), pp. 466–474. DOI: `http://dx.doi.org/10.1016/j.chembiol.2013.03.012`.

[69] B. J. Cafferty and N. V. Hud. “Was a Pyrimidine-Pyrimidine Base Pair the Ancestor of Watson-Crick Base Pairs? Insights from a Systematic Approach to the Origin of RNA”. In: *Isr. J. Chem.* 55 (2015), pp. 891–905. DOI: `DOI:10.1002/ijch.201400206`.

[70] M. J. van Vliet, J. Visscher, and A. W. Schwartz. “An Achiral (Oligo)Nucleotide Analog”. In: *J. Mol. Evol.* 38 (1994), pp. 438–442. DOI: `DOI:10.1007/BF00178843`.

[71] C. V. Mungi et al. “Synthesis of Barbituric Acid Containing Nucleotides and Their Implications for the Origin of Primitive Informational Polymers”. In: *Phys. Chem. Chem. Phys.* 18 (2016), pp. 20144–20152. DOI: `DOI:10.1039/C6CP00686H`.

[72] Y. Valadbeigi, H. Farrokhpour, and M. Tabrizchi. “Theoretical Study on the Isomerization and Tautomerism in Barbituric Acid.” In: *Struct. Chem.* 25 (2014), pp. 1805–1810. DOI: `DOI:10.1007/s11224-014-0452-0`.

[73] J. D. WATSON and F. H. C. CRICK. “Genetical Implications of the Structure of Deoxyribonucleic Acid”. In: *Nature* 171 (1953), p. 964.

[74] M. D. Topal and J. R. Fresco. “Complementary Base Pairing and the Origin of Substitution Mutations”. In: *Nature* 263 (1976), p. 285.

[75] K. Bebenek, L. C. Pedersen, and T. A. Kunkel. “Replication Infidelity via a Mismatch with Watson-Crick Geometry”. In: *Proc. Natl. Acad. Sci.* 108 (2011), 1862 LP–1867.

[76] W. Wang, H. W. Hellinga, and L. S. Beese. “Structural Evidence for the Rare Tautomer Hypothesis of Spontaneous Mutagenesis”. In: *Proc. Natl. Acad. Sci.* 108 (2011), pp. 17644–17648.

[77] D. Shugar and B. Kierdaszuk. “New Light on Tautomerism of Purines and Pyrimidines and Its Biological and Genetic Implications”. In: *J. Biosci.* 8 (1985), pp. 657–668. DOI: `DOI10.1007/BF02702764`.

[78] J. M. E. Vipender Singh Bogdan I. Fedeles. “Role of Tautomerism in RNA Biochemistry”. In: *RNA* 21 (2015), pp. 1–13.

[79] V. Singh et al. “Direct Observation of Multiple Tautomers of Oxythiamine and Their Recognition by the Thiamine Pyrophosphate Riboswitch”. In: *ACS Chem. Biol.* 9 (2014), pp. 227–236. DOI: `DOI:10.1021/cb400581f`.

[80] S. Thore, C. Frick, and N. Ban. “Structural Basis of Thiamine Pyrophosphate Analogues Binding to the Eukaryotic Riboswitch”. In: *J. Am. Chem. Soc.* 130 (2008), pp. 8116–8117. DOI: `DOI:10.1021/ja801708e`.

[81] S. D. Gilbert et al. “Adaptive Ligand Binding by the Purine Riboswitch in the Recognition of Guanine and Adenine Analogs”. In: *Structure* 17 (2009), pp. 857–868. DOI: `DOI:10.1016/j.str.2009.04.009`.

[82] R. Kakkar and V. Katoch. “AM1 Study of Proton-Transfer Reactions of Barbituric Acid”. In: *Int. J. Quantum Chem.* 74 (1999), pp. 327–336.

[83] K. Senthilkumar and P. Kolandaivel. “Quantum Chemical Studies on Tautomerism of Barbituric Acid in Gas Phase”. In: *Journal of Computer-Aided Molecular Design* 16 (2002), pp. 263–272.

[84] S. Ralhan and N. K. Ray. “Density Functional Study of Barbituric Acid and Its Tautomers”. In: *J. Mol. Struct. THEOCHEM* 638 (2003), pp. 83–88. DOI: `DOI:10.1016/S0166-1280(03)00260-4`.

[85] V. B. Delchev. “DFT Ab Initio Study of the Keto-Enol Tautomerism of Barbituric Acid”. In: *J. Struct. Chem.* 45 (2004), pp. 570–578. DOI: `DOI:10.1007/s10947-005-0031-8`.

[86] M. U. Schmidt et al. “The Thermodynamically Stable Form of Solid Barbituric Acid: The Enol Tautomer”. In: *Angew. Chemie - Int. Ed.* 50 (2011), pp. 7924–7926. DOI: `DOI:10.1002/anie.201101040`.

[87] M. G. Marshall, V. Lopez-Diaz, and B. S. Hudson. “Single-Crystal X-Ray Diffraction Structure of the Stable Enol Tautomer Polymorph of Barbituric Acid at 224

and 95K". In: *Angew. Chemie - Int. Ed.* 55 (2016), pp. 1309–1312. DOI: `DOI: 10.1002/anie.201508078`.

[88] Millefiori Arcangelo Millefiori Salvatore. "Tautomerism in Barbituric and Thiobarbituric Acid". In: *J. Heterocycl. Chem.* 26 (1989), p. 639.

[89] V. I. Danilov et al. "On the Mechanism of the Mutagenic Action of 5-Bromouracil: A DFT Study of Uracil and 5-Bromouracil in a Water Cluster". In: *J. Phys. Chem. A* 113 (2009), pp. 2233–2235. DOI: `DOI:10.1021/jp811007j`.

[90] S. A. Trygubenko et al. "Correlated Ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment and in Aqueous Solution Part 1. Cytosine". In: *Phys. Chem. Chem. Phys.* 4 (2002), pp. 4192–4203. DOI: `DOI:10.1039/B202156K`.

[91] Hanus et. al. "Correlated Ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment and in Aqueous Solution. Guanine: Surprising Stabilization of Rare Tautomers in Aqueous Solution". In: *J. Am. Chem. Soc.* 105 (2003), pp. 7678–7688. DOI: `DOI:10.1021/ja034245y`.

[92] M. Hanus et al. "Correlated Ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment, and in Aqueous Solution. Part 3. Adenine". In: *J. Phys. Chem. B* 108 (2004), pp. 2087–2097. DOI: `DOI:10.1021/jp036090m`.

[93] J. Rejnek et al. "Correlated Ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment and in Aqueous Solution. Part 4. Uracil and Thymine". In: *Phys. Chem. Chem. Phys.* 7 (2005), pp. 2006–2017. DOI: `DOI:10.1039/B501499A`.

[94] J. L. Llanio-Trujillo, J. M. C. Marques, and F. B. Pereira. "An Evolutionary Algorithm for the Global Optimization of Molecular Clusters: Application to

Water, Benzene, and Benzene Cation". In: *The Journal of Physical Chemistry A* 115.11 (2011), pp. 2130–2138. DOI: `10.1021/jp1117695`.

[95] Sven Heiles and Roy L. Johnston. "Global optimization of clusters using electronic structure methods". In: *International Journal of Quantum Chemistry* 113.18 (2013), pp. 2091–2109. DOI: `10.1002/qua.24462`.

[96] M. A. Morsy, A. M. Al-Somali, and A. Suwaiyan. "Fluorescence of Thymine Tautomers at Room Temperature in Aqueous Solutions". In: *J. Phys. Chem. B* 103 (1999), pp. 11205–11210. DOI: `DOI:10.1021/jp990858e`.

[97] A. Suwaiyan, M. A. Morsy, and K. A. Odah. "Room Temperature Fluorescence of 5-Chlorouracil Tautomers". In: *Chem. Phys. Lett.* 237 (1995), pp. 349–355. DOI: `DOI:https://doi.org/10.1016/0009-2614(95)00294-E`.

[98] F. Buyl et al. "Matrix-Isolation FT-IR Studies and Ab-Initio Calculations of Hydrogen-Bonded Complexes of Molecules Modeling Cytosine or Isocytosine Tautomers. 3. Complexes of 4-Hydroxypyridine and 3-Hydroxypyridine with H2O in Ar Matrixes." In: *J. Phys. Chem.* 99 (1995), pp. 14967–14979. DOI: `DOI : 10.1021/j100041a010`.

[99] A.D. Beke. In: *Phys. Rev. A* ().

[100] C. Lee, W. Yang, and R.G. Parr. In: *Phys. Rev. B* ().

[101] M. J. Frisch et al. *Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT.* 2009.

[102] Chris Oostenbrink et al. "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6". In: *Journal of Computational Chemistry* 25.13 (2004), pp. 1656–1676. DOI: `10.1002/jcc.20090`.

[103] Alpeshkumar K Malde et al. “An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0”. In: *Journal of Chemical Theory and Computation* 7.12 (2011), pp. 4026–4037. ISSN: 1549-9618. DOI: `10.1021/ct200196m`.

[104] W. H. Press et al. *Numerical Recipes: The Art of Scientific Computing.* 3rd ed. Cambridge University Press: U.K, 2007, p. 1256.

[105] H J C Berendsen, J R Grigera, and T P Straatsma. “The missing term in effective pair potentials”. In: *The Journal of Physical Chemistry* 91.24 (1987), pp. 6269–6271. ISSN: 0022-3654. DOI: `10.1021/j100308a038`.

[106] Shichi Nosé. “A molecular dynamics method for simulations in the canonical ensemble”. In: *Molecular Physics* 52.2 (1984), pp. 255–268. ISSN: 0026-8976. DOI: `10.1080/00268978400101201`.

[107] William G Hoover. “Canonical dynamics: Equilibrium phase-space distributions”. In: *Phys. Rev. A* 31.3 (1985), pp. 1695–1697. DOI: `10.1103/PhysRevA.31.1695`.

[108] M Parrinello and A Rahman. “Polymorphic transitions in single crystals: A new molecular dynamics method”. In: *Journal of Applied Physics* 52.12 (1981), pp. 7182–7190. ISSN: 0021-8979. DOI: `10.1063/1.328693`.

[109] Tom Darden, Darrin York, and Lee Pedersen. “Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems”. In: *The Journal of Chemical Physics* 98.12 (1993), pp. 10089–10092. ISSN: 0021-9606. DOI: `10.1063/1.464397`.

[110] Sander Pronk et al. “GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit”. In: *Bioinformatics* 29.7 (2013), pp. 845–854. ISSN: 1367-4803.

[111] M. M. Brister, M. Pollum, and C. E. Crespo-Hernndez. “Photochemical etiology of promising ancestors of the RNA nucleobases”. In: *Phys. Chem. Chem. Phys.* 18 (30 2016), pp. 20097–20103. DOI: `10.1039/C6CP00639F`.

[112] Jean-Marc L. Pecourt, Jorge Peon, and Bern Kohler. “Ultrafast Internal Conversion of Electronically Excited RNA and DNA Nucleosides in Water”. In: *Journal of the American Chemical Society* 122 (2000), pp. 9348–9349. DOI: `10.1021/ja0021520`.

[113] Michel Mons et al. “Near-UV Resonant Two-Photon Ionization Spectroscopy of Gas Phase Guanine: Evidence for the Observation of Three Rare Tautomers”. In: *The Journal of Physical Chemistry A* 110 (2006), pp. 10921–10924. DOI: `10.1021/jp063738x`.

[114] Christel M. Marian. “The Guanine Tautomer Puzzle: Quantum Chemical Investigation of Ground and Excited States”. In: *The Journal of Physical Chemistry A* 111.8 (2007), pp. 1545–1553. DOI: `10.1021/jp068620v`.

[115] Anders Broo. “A Theoretical Investigation of the Physical Reason for the Very Different Luminescence Properties of the Two Isomers Adenine and 2-Aminopurine”. In: *The Journal of Physical Chemistry A* 102 (1998), pp. 526–531. DOI: `10.1021/jp9713625`.

[116] Eyal Nir et al. “On the Photochemistry of Purine Nucleobases”. In: *The Journal of Physical Chemistry A* 105 (2001), pp. 5106–5110. DOI: `10.1021/jp0030645`.

[117] Kurt A. Kistler and Spiridoula Matsika. “Radiationless Decay Mechanism of Cytosine: An Ab Initio Study with Comparisons to the Fluorescent Analogue 5-Methyl-2-pyrimidinone”. In: *The Journal of Physical Chemistry A* 111 (2007), pp. 2650–2661. DOI: `10.1021/jp0663661`.

[118] Elizabeth Mburu and Spiridoula Matsika. “An Ab Initio Study of Substituent Effects on the Excited States of Purine Derivatives”. In: *The Journal of Physical Chemistry A* 112 (2008), pp. 12485–12491. DOI: `10.1021/jp807145c`.

[119] Nicholas V. Hud et al. “The Origin of RNA and My Grandfather’s Axe”. In: *Chemistry and Biology* 20 (2013), pp. 466 –474. DOI: `doi.org/10.1016/j.chembiol.2013.03.012`.

[120] Dana Nachtigallov et al. “The effect of C5 substitution on the photochemistry of uracil”. In: *Phys. Chem. Chem. Phys.* 12 (19 2010), pp. 4924–4933. DOI: `10.1039/B925803P`.

[121] Zhongneng Zhou et al. “Ultrafast Excited-State Dynamics of Cytosine Aza-Derivative and Analogues”. In: *The Journal of Physical Chemistry A* 121.14 (2017), pp. 2780–2789. DOI: `10.1021/acs.jpca.6b12290`.

[122] Yuyuan Zhang et al. “Excited-State Dynamics of Melamine and Its Lysine Derivative Investigated by Femtosecond Transient Absorption Spectroscopy”. In: *Molecules* 21 (2016), p. 1645.

[123] Delchev Vassil B. and Ivanova Iva P. “Theoretical study of the excited-state reaction paths of the OH and NH dissociation processes in barbituric acid”. In: *Monatsh Chem* 143 (2012), pp. 1141–1150.

[124] Shterev Ivan G. and Delchev Vassil B. “Solvent influence on the excited states of the oxo form of barbituric acid and the mechanisms of the out-of-plane non-radiative elongation of the NAH bond: A comparative theoretical and experimental study”. In: *Spectrochimica Acta Part A* 125 (2014), pp. 384–390.

[125] Rudiger Bauernschmitt and Reinhart Ahlrichs. “Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory”. In: *Chemical Physics Letters* 256.4 (1996), pp. 454–464. DOI: `https://doi.org/10.1016/0009-2614(96)00440-X`.

[126] Spiridoula Matsika. “Radiationless Decay of Excited States of Uracil through Conical Intersections”. In: *The Journal of Physical Chemistry A* 108.37 (2004), pp. 7584–7590. DOI: `10.1021/jp048284n`.

[127] Thomas Gustavsson et al. "Solvent Effects on the Steady-state Absorption and Fluorescence Spectra of Uracil, Thymine and 5-Fluorouracil". In: *Photochemistry and Photobiology* 83.3 (), pp. 595–599. DOI: `10.1111/j.1751-1097.2007.00052.x`.

[128] A Reuther et al. "Femtosecond photo-ionization of nucleic acid bases: electronic lifetimes and electron yields". In: *Chemical Physics Letters* 325.4 (2000), pp. 360–368. ISSN: 0009-2614. DOI: `https://doi.org/10.1016/S0009-2614(00)00699-0`.

[129] Cllia Canuel et al. "Excited states dynamics of DNA and RNA bases: Characterization of a stepwise deactivation pathway in the gas phase". In: *The Journal of Chemical Physics* 122.7 (2005), p. 074316. DOI: `10.1063/1.1850469`.

[130] M. Kotur et al. "Following Ultrafast Radiationless Relaxation Dynamics With Strong Field Dissociative Ionization: A Comparison Between Adenine, Uracil, and Cytosine". In: *IEEE Journal of Selected Topics in Quantum Electronics* 18.1 (2012), pp. 187–194. DOI: `10.1109/JSTQE.2011.2107892`.

[131] Spiridoula Matsika et al. "Ultrafast Relaxation Dynamics of Uracil Probed via Strong Field Dissociative Ionization". In: *The Journal of Physical Chemistry A* 117.48 (2013), pp. 12796–12801. DOI: `10.1021/jp408073d`.

[132] D. S. Wishart et al. "1H, 13C and 15N Chemical Shift Referencing in Biomolecular NMR". In: *J. Biomol. NMR* 6 (1995), pp. 135–140. DOI: `DOI:10.1007/BF00211777`.

[133] Ulrich et. al. "BioMagResBank". In: *Nucleic Acids Res.* 36 (2008), pp. D402–D408.

Appendix A

We found in our study of DHI that the oscillator strengths (f) were close to 0 for $\pi\sigma^*$ states. However, this cannot be explained using symmetry arguments, when we take into account that DHI has C_s symmetry. Therefore, we have looked into the molecular orbitals (MOs) or DHI to explain why the f between the ground (A') and the $\pi\sigma^*$ state (A''), for a molecule having C_s symmetry, (in this case DHI) is close to zero. For reference we have also looked at the f of the $\pi\pi^*$ state as it was found to have a high f. Figure 1 shows the equillibrium structure of DHI. We see that the molecule lies in the Y-Z plane with the X-axis is perpendicular to the plane.

The transition dipole moments calculated using EOM-CCSD method are given in Table 1 and the corresponding orbitals and their products are shown in Figure

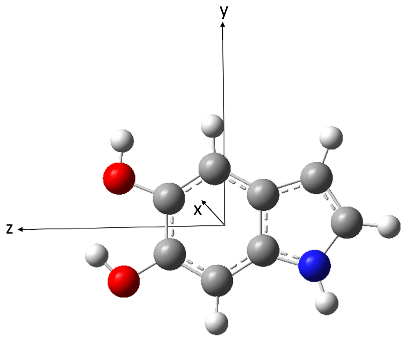


Figure 1: Equilibrium structure of DHI indicated with the Cartesian axes.

Table 1: Transition dipole moments calculated using EOM-CCSD method

State	x	y	z
$S_1\pi\pi^*$	0.000	0.056	0.972
$S_2\pi\sigma^*$	0.123	0.000	0.000

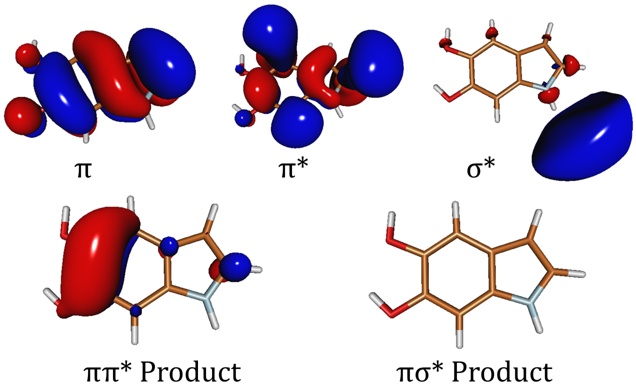


Figure 2: Orbitals involved in S_1 and S_2 state with their products shown below.

2.

We see from Figure 2 that the product of the $\pi\pi^*$ orbitals has a significant value, which can be clearly visualized, however the product of the $\pi\sigma^*$ orbitals is 0. This is intuitve when we see that the σ^* orbital is far from the molecule and does not have any overlap with the π orbital. We wanted to further investigate the transition dipole moment integrals along the three directions, x, y and z.

Orbitals cross-sections using cube files

Cube files are written to provide values for a physical quantity, along a 3D-grid, inside a cubical box containing molecule. The physical quantitites can be electron density, molecular orbital (*viz.* one electron wavefunction). The purpose of this utility is to allow plotting of orbitals (or other physical quantity) by external programs.

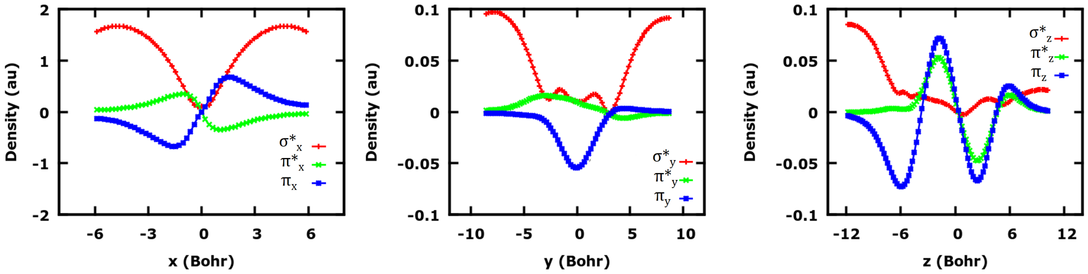


Figure 3: Plots of the leading orbitals taking part in the transitions of the $\pi\pi^*$ and $\pi\sigma^*$ states. They are plotted along the x, y and z axes. The states are appropriately scaled wherever necessary.

Here molecular orbital is written as a scalar field (i.e., $V(x, y, z) = V$). We project out the relevant component of the orbital by integrating the function in perpendicular directions at each point in the given direction.

Shown in Figure 3 are the plots of the leading orbitals taking part in the transitions of the $\pi\pi^*$ and $\pi\sigma^*$ states, i.e. lowest–σ^*, highest–π, and lowest–π^* orbitals, respectively. These plots represent the corresponding orbitals plotted along the x, y and z axes, respectively, by integrating out the other two directions.

Figure 3 can be read as the value of the function at a particular value of x, i.e., the isosurface value. We see that the plots representing the π and π^* orbitals have a high value close to $x = 0$, which indicates the region close to the molecule, and values close to 0 at large x. The opposite is true for the σ^* orbital. This is because the π and π^* orbitals are mostly localised over the molecule, while the σ^* orbital is largely present outside the molecule. Figure 4 shows the product of the $\pi\pi^*$ and $\pi\sigma^*$ orbitals.

From these plots in Figure 4, we can 'weakly' rationaliz the transition dipole moment values obtained from EOM-CCSD calculation. The calculation for the same with values obtained from cube file yields the result presented below in Table 2

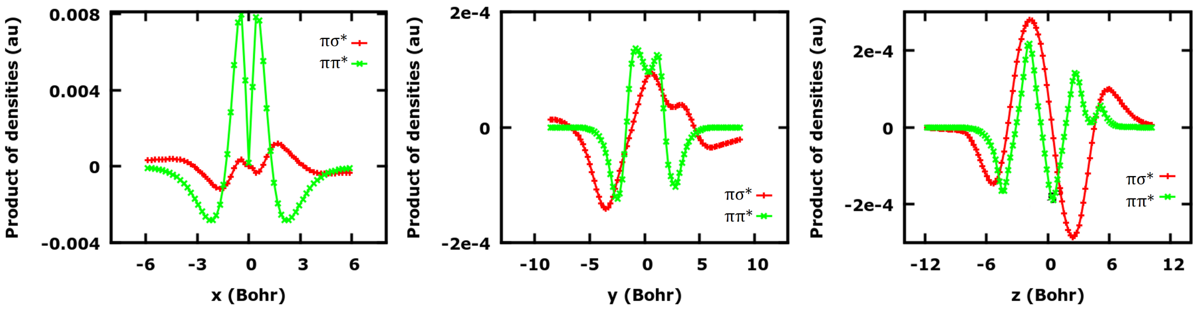


Figure 4: Plots of the product of the leading orbitals taking part in the transitions of the $\pi\pi^*$ and $\pi\sigma^*$ states. They are plotted along the x, y and z axes. The states are appropriately scaled wherever necessary.

Table 2: Transition dipole moments calculated using the overlap integrals between molecular orbitals.

State	x	y	z
$S_1\pi\pi^*$	1.60×10^{-5}	0.0271	0.749
$S_2\pi\sigma^*$	2.30×10^{-3}	8.82×10^{-4}	9.14×10^{-4}

Appendix B

The IR experiments were performed by Mr. Pranab Deb and Dr. Sayan Bagchi, CSIR-NCL Pune.

Experimental details: IR spectroscopy

IR absorption spectra were recorded on a Vertex 70V FTIR Spectrometer (Bruker) with 2cm^{-1} resolution at room temperature. For each solution, 60Lof the sample solution (25 mM) was loaded into a demountable cell consisting of two CaF2windows (3 mm thickness) sepa-rated by a mylar spacer of 56 m thickness. Solvent IR spectra were also acquired, under identical conditions, and used to subtract the background. An inbuilt polnomial fitting algorithm available with the instrument was used to baseline each solvent subtracted FTIR spectrum.

Studying Solvation with IR Spectroscopy

We have investigated the solvent effect on the keto-enol tautomerization of BA using IR spectroscopy. IR spectra of BA has been recorded in acetonitrile, D_2O (water) and 1:1 mixture (v/v) of acetonitrile:water (Figure 5)and analysed for the frequency range of 1500-1800 cm^{-1}. This frequency range was chosen to monitor both C=O and C=C stretch IR signals in solution at different water content.

In neat acetonitrile, BA shows two distinct peaks at 1724 cm^{-1} and 1743 cm^{-1}. However, no signal from the C=C stretch is observed (1600 cm^{-1}) in the absence of water. The peak at 1724 cm^{-1} has been assigned to the C=O stretch while the peak at 1743cm^{-1} arises from Fermi resonance (see SI for more details). The observation of a single C=O peak in spite of the presence of three C=O groups in BA indicates similar environment for the C=O groups in BA. In 1:1 (v/v) mixture of acetonitrile and water, the C=O peak is red shifted from acetonitrile to 1697 cm^{-1} and a new peak at 1588 cm^{-1} arises. The appearance of the peak at 1588 cm^{-1} indicates the presence of a C=C (enol form) in an aqueous environment. However, the complete absence of this peak in neat acetonitrile demonstrates that the keto form is the dominating tautomer in the absence of water. In neat D_2O, the C=O peak is observed at 1693cm^{-1} and the C=C absorbs at 1587 cm^{-1}. Moreover, the intensity of the C=C is greater in neat water, as compared to the acetonitrile-water solution. The enhancement in intensity of the C=C peak reveals the increased stability of the enol isomer in an aqueous environment. To confirm that the enol is coming from the C_5 hydrogens not from the N_1 and N_3 hydrogens, we have performed the controlled IR experiment of methyl substituted BA (DMBA) in acetonitrile and D_2O (Figure 6). The IR spectra of DMBA are very much similar to the BA. The observation of the C=C peak at 1597 cm^{-1} in case of DMBA further supports that the double bond is arising from the C_5 hydrogens of BA. The appearance of the C=C peak in an aqueous environment and its enhanced intensity with increasing water context corroborates the theoretical prediction of water mediated stability of the enol form of BA.

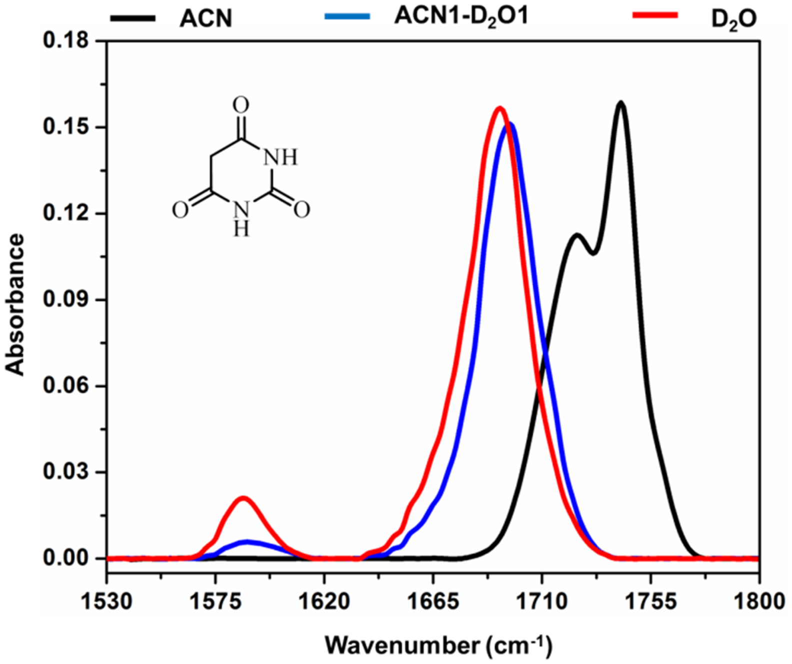


Figure 5: FTIR spectra BA in acetonitrile (black), D2O (red) and 1:1 acetonitrile- D2O mixture (blue).

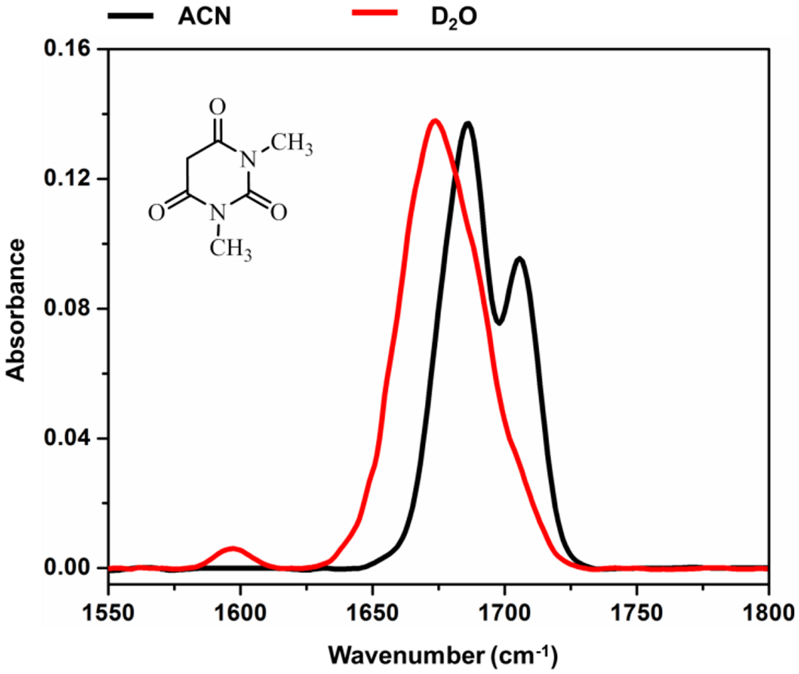


Figure 6: FTIR spectra BA in acetonitrile (black), D2O (red) and 1:1 acetonitrile- D2O mixture (blue).

Appendix C

The NMR experiments were performed by Mr. Harshad Paithankar and Dr. Jeetender Chugh, IISER Pune.

Experimental Details: NMR Spectroscopy

All the NMR experiments were carried out on a Bruker AVANCE III 600 MHz NMR spectrometer equipped with a quad-channel (1H/13C/15N/31P-2H) cryogenic probe at a constant temperature of 298 K. For NMR data recording, 100 M solution of barbituric acid was prepared in (1) 50 mM Sodium phosphate buffer pH 7.4 (with 10% D2O), and in (2) deuterated acetonitrile (ACN-D3). The 13C 1D-NMR (with proton decoupling during acquisition) spectra for the two samples were acquired using a 30 flip angle. For DEPT-13535 experiments (with proton decoupling during acquisition), the polarization transfer from proton to carbon was achieved using a shaped pulse (2 ms composite smoothed Chirp pulse with 4000 points), with 135 read pulse to distinguish CH2 from CH and CH3 carbons. In both the experiments for the two samples, data was recorded with 512 scans, 64k data points in a spectral width of 238.8 ppm and an inter-scan relaxation delay of 2 s. Carbon chemical shifts were indirectly referenced to DSS as described elsewhere.[132]

NMR Spectroscopy to Characterize Isomeric Form

The 13C 1D-NMR spectrum recorded in sodium phosphate buffer pH 7.5 showed six distinct peaks corresponding to four carbons in BA (Figure 7a) where a set of three peaks were observed as high intensity peaks when compared with the other set of three peaks. This suggested presence of two species with different populations present in slow exchange in this buffer condition. The high intensity peaks at δ 80.8, δ 156.7 and δ 171.1 ppm were assigned to the carbons C_5, C_2 and C_4/C_6, respectively, marked with orange labels in Figure 7a and 7c. This assignment, when compared with literature[133], suggested for the presence of enol tautomer of BA (Figure 7c). The set of three small intensity peaks (δ 155.4, δ 172.5 and δ 41.4 ppm assigned to C_2, C_4/C_6 and C_5, respectively, marked with green labels in Figure 7a and 7c) suggested the presence of a small population (20%) of keto tautomer of BA in sodium phosphate buffer. The DEPT-135 showed no peaks at δ 156.7 and δ 171.1 ppm confirming the corresponding carbons (C_2 and C_4/C_6) as quaternary carbons. A positive peak in DEPT-135 at δ 80.8 ppm confirmed the presence of CH2–carbon (C_5)corroborating the presence of enol tautomer. C_5 in keto form (δ at 41.4 ppm) did not show any negative peak (as expected for a –CH2– group) in DEPT-135 in sodium phosphate buffer due to very low s/n. The 13C 1D-NMR spectrum of BA, recorded in ACN-D3, showed three peaks (Figure 7). The peaks at δ 41.6, δ 153.1 and δ 169.3 ppm were assigned to C_5, C_2 and C_4/C_6, respectively. The peak at δ 41.6 ppm showed a negative peak in DEPT-135 suggesting –CH2– carbon represented by C_5 in keto tautomer of BA. Thus, while in aqueous buffer, BA is found to be present as a mixture of keto and enol forms, it existed in pure keto form in absence of water.

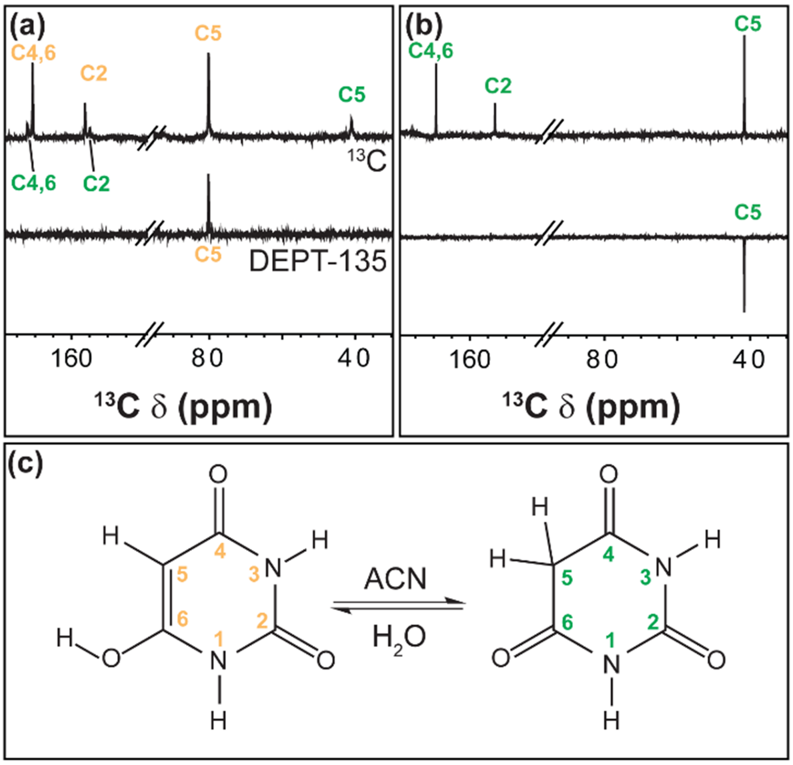


Figure 7: 13C 1D-NMR and DEPT-135 spectra of BA in (a) 50 mM sodium phosphate buffer pH 7.4 (with 10% D2O), (b) ACN-D3. (c) Switching of tautomeric stability of BA in ACN and water.

Appendix D

To study the effect of solvation, quantum optimization of the solute surrounded by water molecules is an excellent method. Quantum optimizations are very sensitive to the starting geometry and to adequately sample all energetically close minima, one needs to start with all possible structures of the solute and water molecules. However, with increasing number of water molecules, this becomes a cumbersome task. Here, we present a protocol, which uses Molecular Dynamics (MD) simulations to choose such clusters while studying the effect of solvation.

1. Perform an MD simulation solvating the solute molecule. In our study, we performed MD simulations for both the tautomers separately.

2. Extract the distribution of hydrogen bonds between solute and solvent molecules (Figure 4.4).

3. If the solute has multiple centres for forming hydrogen bonds with the solvent, then perform additional micro-clustering analysis by determining most probable ways by which the total number of solute-solvent hydrogen bonds are formed (Figures 4.5, 4.6, 4.7 and 4.8).

4. Extract the coordinates of solute and desired number of water molecules from any 10 MD frames, which belong to the most probable distribution of hydrogen bonds, as discussed in Steps 2 and 3 above. Use these extracted

geometries as starting structures for quantum optimization.

5. In order to extract the geometries as mentioned in Step 4, we need to remove the unwanted/excess water molecules, which are not directly hydrogen bonded to the solute. This leads to an abrupt alteration of the hydrogen bond network already present between the water molecules in the MD frame. Therefore, it is possible that when this frame is quantum optimized, the number of solute-solvent hydrogen bonds, and/or distribution of hydrogen bonds at each centre, may change with respect to the initially chosen MD frame. In such cases, discard the structure and repeat Step 4 to choose another frame satisfying the desired conditions, and perform the quantum optimization once again.